

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2001, 02:50:15 ; Search time 1697.44 Seconds
(without alignments)
4205.899 Million cell updates/sec

Title: US-09-388-090-3
Perfect score: 1395
Sequence: 1 gtgttcaaaaataccaata.....tcagggaagtcacctcaa 1395

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 1118133 seqs, 2558875100 residues 2236266
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba1:
2: gb_ba2:
3: gb_om:
4: gb_ov:
5: gb_ph:
6: gb_pl1:
7: gb_pl2:
8: gb_pri:
9: gb_pr2:
10: gb_pr3:
11: gb_ro:
12: gb_sy:
13: gb_un:
14: em_fun:
15: em_hum1:
16: em_hum2:
17: em_in:
18: em_om:
19: em_or:
20: em_ov:
21: em_pat:
22: em_ph:
23: em_pl:
24: em_ro:
25: em_sts:
26: em_sy:
27: em_un:
28: em_vi:
29: gb_ba3:
30: gb_in1:
31: gb_in2:
32: gb_in3:
33: gb_pl3:
34: gb_pr4:
35: em_ba1:
36: em_ba2:
37: em_htg1:
38: em_htg2:
39: em_htg3:
40: em_htg4:
41: em_htg5:
42: em_htg6:
43: em_htg7:

44: em_htg8:
45: em_htg9:
46: em_htg10:
47: em_hum3:
48: em_hum4:
49: em_hum5:
50: em_hum6:
51: gb_pr5:
52: gb_pr6:
53: gb_pr7:
54: gb_htg1:
55: gb_htg2:
56: gb_htg3:
57: gb_htg4:
58: gb_htg5:
59: gb_htg6:
60: gb_htg7:
61: gb_htg8:
62: gb_htg9:
63: gb_htg10:
64: gb_htg11:
65: gb_htg12:
66: gb_htg13:
67: gb_htg14:
68: gb_htg15:
69: gb_htg16:
70: gb_htg17:
71: gb_htg18:
72: gb_htg19:
73: gb_htg20:
74: gb_htg21:
75: gb_htg22:
76: gb_htg23:
77: gb_sts1:
78: gb_sts2:
79: gb_vil:
80: gb_vil2:
81: gb_pat1:
82: gb_pat2:
83: em_htg0:
84: gb_htg24:
85: gb_pr8:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	1322.2	94.8	10531	1	AE002409	Neisseria
c 2	1320.6	94.7	1500	2	AX011576	Sequence
c 3	1315.8	94.3	311321	2	NMA322491	AX011576 Sequence
4	1263	90.5	1500	2	AX011574	Sequence
5	1258.2	90.2	1500	2	AX011572	Sequence
6	1061	76.1	1110	2	AX011578	Sequence
7	365.6	26.2	1961	2	PAU32853	Sequence
8	365.6	26.2	4587	2	PAU49151	Pseudomonas
9	365.6	26.2	10977	1	AE004511	Pseudomonas
10	274.8	19.7	2027	2	AV030799	Pseudomonas
11	260.4	18.7	12245	1	AE004037	Azotobacter
12	260.2	18.7	390	2	AX011580	Xylella f
13	246	17.6	1362	2	AF293977	Sequence
14	217.4	15.6	11461	1	AE004859	Aeromonas
15	212.4	15.2	1974	2	RM031512	Pseudomonas
16	211.8	15.2	2009	2	BAU07352	Rhizobium m
17	197	14.1	1905	2	BRUHTRAH	Brucella ab
c 18	196.6	14.1	10566	1	AE004142	L09274 Brucella ab
19	196	14.1	2714	2	ECU32495	AE004142 Vibrio ch
20	196	14.1	3156	2	ECU15661	Escherichia
21	196	14.1	10713	1	AE000402	Escherich

22	196	14.1	110000	2	EC0UW67.1	Continuation (2 of
23	186.6	13.4	1191	2	PAU29172	U29172 Pseudomonas
c	24	183.4	13.1	18424	1	AF163841 Myxococcus
	25	181.2	13.0	1476	1	AB040030 Shigella
	26	178	12.8	10199	1	AE000125 Escherich
	27	178	12.8	82727	2	EC082K
	28	178	12.8	123171	2	ECU70214
	29	176.4	12.6	1854	81	AR095624
	30	176.4	12.6	1855	2	EC0HTRAA
	31	176.4	12.6	1855	2	EC0HTRAA
	32	175.6	12.6	1980	29	STSPHSPG
	33	175.6	12.6	1980	81	AR037634
	34	175.6	12.6	1980	81	AR084295
	35	175.2	12.6	1980	81	AI8802
	36	164.6	11.8	4151	2	RCFDXE
	37	164.4	11.8	1800	2	D78376
	38	157.2	11.3	1613	2	BAU07351
	39	154.2	11.1	1791	81	AR000272
	40	154.2	11.1	1791	81	I51141
	41	154.2	11.1	1791	82	I78681
	42	154.2	11.1	1889	2	ROCHTRA
	43	149.6	10.7	11201	1	AE001732
	44	148.2	10.6	11881	1	AE001355
	45	147.6	10.6	2778	2	BJY13616

ALIGNMENTS

RESULT	1					
AE002409/c						
LOCUS	AE002409	10531 bp	DNA	BCT	25-MAY-2000	
DEFINITION	Neisseria meningitidis serogroup B strain MC58 section 51 of 206 of the complete genome.					
ACCESSION	AE002409 AE002098					
VERSION	AE002409.1 GI:7225757					
KEYWORDS	Neisseria meningitidis MC58.					
SOURCE	Neisseria meningitidis MC58.					
ORGANISM	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.					

REFERENCE	1 (bases 1 to 10531)					
AUTHORS	Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C., Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F., Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D., Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D., Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E., Citterone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H., Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V., Pizzi,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R., Rappuoli,R. and Venter,J.C.					
TITLE	Complete genome sequence of Neisseria meningitidis serogroup B strain MC58					
JOURNAL	Science 287 (5459), 1809-1815 (2000)					
MEDLINE	20175755					
PUBMED	10710307					
REFERENCE	2 (bases 1 to 10531)					
AUTHORS	Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C., Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F., Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D., Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D., Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E., Citterone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H., Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V., Pizzi,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R., Rappuoli,R. and Venter,J.C.					

TITLE	Direct Submission					
JOURNAL	Submitted (17-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA					
FEATURES	Location/Qualifiers					
source	1..10531					
	/organism="Neisseria meningitidis MC58"					
	/strain="MC58"					

gene	/db_xref="taxon:122586"	
	/note="serogroup: B"	
	complement(132..1631)	
	/gene="NMB0532"	
	/note="This region contains an authentic frame shift and is not the result of a sequencing artifact; similar to PID:1184884 PID:1220194 percent identity: 60.59; identified by sequence similarity; putative:protease DO, authentic frameshift"	
gene	complement(1769..2398)	
	/gene="NMB0533"	
CDS	complement(1769..2398)	
	/gene="NMB0533"	
	/note="similar to GB:J02857 SP:P20625 PID:146972 GB:U00096 PID:1742691 percent identity: 86.12; identified by sequence similarity; putative"	
	/codon_start=1	
	/transl_table=11	
	/product="endonuclease III"	
	/protein_id="AAF40962.1"	
	/db_xref="GI:7225758"	
	/translation="MNRHIREIFERFRAANPHPTTELNFNSPPELLIAVLASAQATD VGNKATAKLPVADTPQAMLDLGLDGVMEYTKTIGLYKTKSKIIIMOTCRLLLEKYN GEPEDREALESLPCVGRKKTANVINTAFCHPVAVDTHIFRVSNRTKIAQCKDREVE DKLMRIPKEFLMDAHHLLHGHRYTCKALKPOCOTCIINDLCYPAKA"	
gene	complement(2444..2857)	
	/gene="NMB0534"	
CDS	complement(2444..2857)	
	/gene="NMB0534"	
	/note="conserved hypothetical protein; identified by Glimmer2; putative"	
	/codon_start=1	
	/transl_table=11	
	/product="conserved hypothetical protein"	
	/protein_id="AAF40963.1"	
	/db_xref="GI:7225759"	
	/translation="MIPSYTYAALAFSAFTSATLLPCTSEAAAFVHRRPEHAYGAL LAGLANGLGSMVSYNMGRLLPSRKMPSEKTLNLMRRFGIWLIAFTLPLVVGDAIPIA AGWLRLNPWTSGLMVLIGTKTARYAFILWGMQYAA"	
gene	3390..4670	
	/gene="NMB0535"	
CDS	3390..4670	
	/gene="NMB0535"	
	/note="similar to PID:1171339 percent identity: 71.89; identified by sequence similarity; putative"	
	/codon_start=1	
	/transl_table=11	
	/product="glucose/galactose transporter"	
	/protein_id="AAF40964.1"	
	/db_xref="GI:7225760"	
	/translation="MSISIFLKRVIHRIHQIMSAQSNHITSPVLVLTLLFFMMGF ITCMNDILIPHLKEIFDLSYQAMLIQCFETAYAVMSIPMGAFVGVKNGVIGGF LITAVGCLLYPAAGSHSYAVFLGALFILASGVTLLOVAGNPYVTLAKPKESATLT LVQAFNALGTTTAPQICAFILADATOTVSKAQOISSVQIPIGLAGLIIILAVFVKM IGLPDARKAAEESAHNHDKTSVQYKHIVFGTAPFCYVGAEVSGISLMNVLYYL KGLDHASAAHLSFYWGAMVGRFLGSAVMAKPAPNRYLAFNAAVVLAVAMATGS GNADVAMSLAIGFFNSIMFPTIFSLATKGLKGTNAASGVLCTAIVGGAVPVVGQ WYADVTYLMSSFFVSVICYLIYVFAVYGYRADK"	
gene	4961..6340	
	/gene="NMB0536"	
CDS	4961..6340	
	/gene="NMB0536"	
	/note="similar to GB:M73530 SP:P27611 PID:143245 PID:2209268 percent identity: 65.57; identified by sequence similarity; putative"	
	/codon_start=1	
	/transl_table=11	
	/product="Na+/H+ antiporter"	
	/protein_id="AAF40965.1"	
	/db_xref="GI:7225761"	
	/translation="MEAFKSLDDMPRGEALAVVVALIAAMGYTIIISLEWLPUMSIIAA IVVLIYGLARGLKYNDMOCQICALNOCMGATYILFFFIGLMVSALMMSCAIPLMYY GFLIISPTVYFSSPALCVIGSVIGSSLLTTCATVGVAFPMGMAAFQADNMTAGAI V	

gene	complement(6469. .7293)
CDS	/gene="NMB0537" complement(6469. .7293) /gene="NMB0537" /note="conserved hypothetical protein; identified by Glimmer2: putative" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAF40966.1" /db_xref="GI:7225762" /translation="MDCAQPKTNFPERLIARAREPDAEDVLNLLRQAEHOFVFDAD TLRLERVLDFSDLEVRDAMITVPEQFHLKSLRSMNVLKENDSERTIYAV IDTAHSRPPYIGEDKD EVLGLHAKDLUKTFMPEQFHLKSLRSMNVLKENDSERTIYAV IDTAHSRPPYIGEDKD EYGTSGLVFEDIEIQVGEIEDEFDEDSADNIHAVSERWRIRHAATEIDINTFF GTEYSSEADTIGGLVTEQLGHLFVRGEKVLIGLQFTVARADNRRLHTLMATRVK" complement(7301. .7810) /gene="NMB0538" complement(7301. .7810) /gene="NMB0538" /note="conserved hypothetical protein; identified by Glimmer2: putative" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAF40967.1" /db_xref="GI:7225763" /translation="MKRTKYPFLTLQRQRLHFNENASSAAGIPAERDFYRWASAL KNEYRDAISILLDEEARAYNDRYRKDYATNVLFSALNEGEIILPCOVSEKLYGDL IICQVVLKRAEAGKTPQEQFAHLTTHGTLHMGVDHIEDDEAEINAEIEIRLMRAA GFPNPYQED" 7929. .8864 /gene="NMB0539" 7929. .8864 /gene="NMB0539" /note="similar to PID:151255 PID:496214 SP:Q60169 percent identity: 72.96; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="porphobilinogen deaminase" /protein_id="AAF40968.1" /db_xref="GI:7225764" /translation="MNPKKLVIASRESLLAMWQAKHIIQGRILKALYPDCEVEIILGMTR GDOIILDKTLKVGKGLFVELEQALYDGRADLAVHS IKDVPMDLPGEFALAAIGBRA NPFDAFVSNQYTRLEEMPEGAVGCTSSLRREAOLRARYPHILLIKPLRGNVOTRLSKLD NGEYDAIILAAAGLQRLQIRIMLISESDSLPAGQGALEIIRAHREDLIEYVLP LNHGVTNACVTAERALARALAGGSCQVPLAAYCTEENGLLRLGLVHPDGVSVVRADA QAPAEYADALGRAVAKLADGDAARELIGALNTEN" complement(9268. .10461) /gene="NMB0540" complement(9268. .10461) /gene="NMB0540" /note="similar to GB:L42023 SP:P44425 PID:1007884 PID:1221762 PID:1205848 percent identity: 84.60; identified by sequence similarity: putative" /codon_start=1 /transl_table=11 /product="aspartate aminotransferase" /protein_id="AAF40969.1" /db_xref="GI:7225765" /translation="MPEKHIIEAAPADPILGIGRAFAETPEKVNILGIGVYKDASGT PLTVAAEKRLLESETTKNYLTIDGVADYNAQOITLLPGKDHEIILASPRAKTAOSL GGTCALRIAEFAKRLQAQNTINISNTPWPHHNAIAKAVGIQDKPTIYDAAKHGLDW DGMTELSQAQGDIVLKHTELLASSYKNGFYNERVGAFTLVAEDEKPLFDFAYQGF NGUEEADYGLRFLKNTTELLASSYKNGFYNERVGAFTLVAEDEKPLFDFAYQGF KTIIRTYNSPASHGANTIALVLKNDOLQAQIAELDEMRGIRKAMROK FVGLLKAAG

```
Qy 961 gccagcggcgattgattgcccataatcttcccgagcccccgagaaacgtgccggcctg 1020
Db 671 GCCGGCGCGCAGCTGATTGCCAAATCTGCGCGCAGCGCCGCGAGAACTGCCGSCCTG 612
Qy 1021 cagcggcgagcatgctcctcagcctcagcggcgagaaatacgtttcttccggcgacctt 1080
Db 611 CAGCGCGCGCAGCATGCTCTCAGCCTCAGCGCGGAGAAATACGTCTTTCGGCGCACCTT 552
Qy 1081 cccgtcatggtcgcgccalacgcgcggaaaaaglcagcctcgcgatagggcga 1140
Db 551 CCCGTTATGTCGGCGCCATTACCGCGGGAAGAGTACAGCCTCGCGGTATGGCGCAAA 492
Qy 1141 ggcgaagaatacacaatacaagcgaagctgggaacgcgcgcgagcatataccggcgcatca 1200
Db 491 GCGGAAGAATAACAATCAAGTCAAGCTGGGCAACGCGCGGAGCATATCGCGCATCA 432
Qy 1201 tccaaacagatgaagccccctacaccgaagaacatccggtaactctcgtcgaatcc 1260
Db 431 TCCAAAACAGATGAAGCCCCCTACACCGAACAGCAATCGGTAACGTTCCTCGTCAATCC 372
Qy 1261 gaggcattaccctcagacacataccagcagcagcgaacacactcgtcgtcgtacgg 1320
Db 371 CGAGCATTACCTTTCAGACACATACCCAGCACGCGCGGAGACCTCTGCTCGTACGG 312
Qy 1321 gttccgacgcggcagaacgcgcggcttaaggcagcgcgcgagaaatccctagcgtcagg 1380
Db 311 GTTTCGAGCGCGGAGACGCGCAGGCTTGAGCGCGCGAGCAAAATCTTGCCGTC -GG 253
Qy 1381 gcaagtcgcccgctcaa 1395
Db 252 GCAAGTCCCGCTCAA 238

RESULT 2
AX011576 LOCUS AX011576 1500 bp DNA BCT 18-SEP-2000
DEFINITION Sequence 5 from Patent WO955872.
ACCESSION AX011576
VERSION AX011576.1 GI:9998107
KEYWORDS
SOURCE
ORGANISM Neisseria meningitidis.
Neisseria meningitidis
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE 1 (bases 1 to 1500)
AUTHORS Ruelle,J.L.
TITLES Basb013 dna and proteins from neisseria meningitidis
JOURNAL Patent: WO 955872-A 04-NOV-1999;
RUELLE JEAN LOUIS (BE); SMITHKLINE BEECHAM BIOLOG (BE)
FEATURES
Location/Qualifiers
source 1..1500
/organism="Neisseria meningitidis"
/db_xref="taxon:487"
BASE COUNT 385 a 471 c 382 g 262 t
ORIGIN

Query Match 94.7%; Score 1320.6; DB 2; Length 1500;
Best Local Similarity 97.1%; Pred. No. 3.4e-234;
Matches 1355; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

Qy 1 glgttcaaaaaataccaataacttcgcttggcgccactgtgtgccccttgcgtgcagcg 60
Db 1 GTGTTCAAAAAATACCAATACCTCGCTTTGGCAGCACTGTGTGCGCCCTCGCTGGCAGGC 60
Qy 61 tgcgaagaaggcagcagcttttgcgtgcgcgacaaaaaagaagcatcctcttagaagcg 120
Db 61 TGCACAAAGGCGAGCGAGCTTTTCGGTTCGGACGACAAAAAAGAGCATCTCTGTAGAAGCG 120
Qy 121 atcgacacacacaaagcagcgcgagtgatgctgctgcccgcacttgcaccaactg 180
Db 121 ATCGAACACACAAAGACGACGCGAGGCTCAGTATGCTTGCTGCCCCGACTTTTGCCCAACTG 180
```

```
Qy 181 gttcaaaagcgaagcccgccggtcgtcgtcaattcaggcagcccccgcccgccgcacc 240
Db 181 GTTCAAAAGTAGGTTCGGCAGTCTGTCNAATTATTCAGGCAGCCCCCGCGCGCACCCNA 240
Qy 241 aacggcagcgccgaatgccgaaacccgaattccgagcccggttgcgacagcgacccgttctac 300
Db 241 AACGGCAGCGGCAATGCCGAAACGATTTCCGACCGGATTTCCGACACGACCGCGTCTTAC 300
Qy 301 gaattttcaaacgcctcgtcccgaaacatgcccgaaatcccccaagaagaagcaatgac 360
Db 301 GAATTTTTCAAAGCCCTTCGTCCCGAAATATGCCCGAAATTCGCCCAAGAAAGACAGATGAC 360
Qy 361 ggcggaatgaactcgttccgggcttcatacatcagcaaaaaacggctacatcctgacccat 420
Db 361 GCGGATTAACCTCGGCTTCGGGCTTCATCATCAGCAAAAGACGGCTACATCTGTACCAAT 420
Qy 421 acccaagtcgttcccggtatgggcagtcataaaagtccctgcataacgacgaagcgcaat 480
Db 421 ACCCAAGTCGTTACCGGCATGGGCAGTATCAAAAGTCCCTGCTCAACGACAAAGCGCAATAT 480
Qy 481 accgcaaacatcagtcgttcgagatgccaaalccgcatgctgcccctctgaaaaatcagcga 540
Db 481 ACCGCCAAACTCATCGGTTCCGATGTCCAAATCCGATGTTCGCCCTTCTGAAATTCAGCGCA 540
Qy 541 acggaagagctaccggtcgtcaaaatcggcaatccccaaaaatttgaaacccggggcaatgg 600
Db 541 ACGGAAGAGCTGCCGCTCGTCNAATCGGCNAATCCCAAGATTTGAAACCGGGCGAATGG 600
Qy 601 gtcgctgccatcgggcgcccttcggcttggacaaacagcgtgacgcgcggcatcgtgtcc 660
Db 601 GTCGCGCCATTCGGCGCGCCCTTCGGCTTCGACAAACAGCGTGACGCGCGGATCTGTGTC 660
Qy 661 gccaaaggcgaagcctgcccaacgaagctacacaccccttcacaaacgaagcttacc 720
Db 661 GCCAAAGGCGAAGCCTGCCCAACGAAGCTACACACCCCTTCATCCAAACGACGCTTACC 720
Qy 721 atcaatccggcgcaattcccgcgcccgctgttcaacttaaaagacagcgtcgtcggcctc 780
Db 721 ATCAATCCGGCAACTCCGCGGCGCCGCTGTTCNACTTAAAGGACAGCTGCTCGGCATC 780
Qy 781 aattcgcaataatacagccgcagcgcggttccttggttcacgtcctcttgcacatcccgatt 840
Db 781 AACTCGCAATATACAGCCGCGCGGCGGATTCATGGCGATTTCTTCGCCATTCGCGATT 840
Qy 841 gacgttgccatgaatgctcgcgcaacagctgaaacaccccgcaagtcacacgcgagca 900
Db 841 GACGTTGCCCATGAATGTGCGCGGAACACAGCTGAAAAAACACCCGCAAGTCCAAACGCGGACAA 900
Qy 901 ctggcggtgattattcagggaagatcctcagcgttggcagcagtcgttcgttctggataaa 960
Db 901 CTGGCGGTGATTATTCAAGAAGTATCTACGGTTTGGCACAACTGTTTCGGCTTGGACAAA 960
Qy 961 gccagcgcgcatgattgccaataatccttcgcggcgcccgacgaacgtgcggcgctg 1020
Db 961 GCCGCGCGCGCACTGATTGCCAAATCTTCGCCGCGAGCCCCCGAGAAAGTTCGCCGCTG 1020
Qy 1021 cagcgggcgacatcgtcctcagcctcgacggcgagaaaaatacagcttctccggcgacctt 1080
Db 1021 CAGCGCGCGCACATCGTTCCTCAGCCTCGCGCGGAAAGAGTACGCTTCGGCGACCTT 1080
Qy 1081 cccgtcatgctcgcgccattacgcccgggaaagaagtcagcctcggcgtatggtgcgaaa 1140
Db 1081 CCGCTTATGTCGGCGCATTAAGCGCGGAAAGAGTACGCTTCGGCGTATGCGGCAAA 1140
Qy 1141 ggcgaagaatacacaatacaagcgaagcgaacgtgggaacgcccgcgagcatcccgcgatcca 1200
Db 1141 GGCGAAGAAATCAAAATCAAAAGTCAAGCTGGGCAACGCCCGCGAGCATATTCGGCGCATCA 1200
Qy 1201 tccaaacagatgaagccccctacaccgaagcaatcccgacgtlccggtcgaatcc 1260
Db 1201 TCCAAAACAGATGAAGCCCCCTACACCGAACAGCAATTCGGTATGCTTGCTGCAATCC 1260
```



```

Query Match          94.3%;      Score 1315.8;  DB 2;      Length 311321;
Best Local Similarity 96.9%;      Pred. No. 8.5e-234;
Matches 1352;      Conservative 0;      Mismatches 42;      Indels 1;      Gaps 1;

```

Qy	1081	cccgtaatggtcgccgcaattacccgcggaagaagtcagcctcggtatggcgcaaa	1141			
Db	9741	cccggttatggtcgccgccaattacgcccggaaaaagaaagtcagcctcggtatggcgcaaa	9682			
Qy	1141	ggcgagaagaatcacatcaaaagcccaagctgggcaacgcccgcgagcatataccggcgcatca	1200			
Db	9681	ggcgagaagaatcacatcaaaagctgggcaacgcccgcgagcatataccggcgcatca	9622			
Qy	1201	tccaaaacagatgaagcccccctacacccgaacagcaatccggtacgttctcggtcggaatcc	1260			
Db	9621	tccaaaacagatgaagcccccctacacccgaacagcaatccggtacgttctcggtcggaatcc	9562			
Qy	1261	gcagcattacccttgaagacacataccagcagcagcggaacacacccctcgctcgtagcgg	1320			
Db	9561	gcagcattacccttgaagacacataccagcagcagcggaacacacccctcgctcgtagcgg	9502			
Qy	1321	gtttccgacgcggcagacgcgcaggcttaagcagcgcgagcaaaatcctagccgtcagg	1380			
Db	9501	gtttccgacgcggcagacgcgcaggcttaagcagcgcgagcaaaatcctagccgtcagg	9443			
Qy	1381	gcaagtcgcccgctcaa	1395			
Db	9442	gcaagtcgcccgctcaa	9428			
RESULT 4						
AX011574						
LOCUS	AX011574	1500 bp	DNA			
DEFINITION	Sequence 3 from Patent WO955872.					
ACCESSION	AX011574					
VERSION	AX011574.1	GI:9998106				
KEYWORDS	Neisseria meningitidis.					
SOURCE	Neisseria meningitidis.					
ORGANISM	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;					
REFERENCE	1 (bases 1 to 1500)					
AUTHORS	Ruelle,J.L.					
TITLE	Bab013 dna and proteins from neisseria meningitidis					
JOURNAL	Patent: WO 955872-A 04-NOV-1999;					
FEATURES	RUELLE JEAN LOUIS (BE); SMITHKLINE BEECHAM BIOLOG (BE)					
source	Location/Qualifiers					
	1..1500					
	/organism="Neisseria meningitidis"					
	/db_xref="taxon:487"					
BASE COUNT	370 a	469 c	392 g			
ORIGIN	269 t					
Query Match 90.5%; Score 1263; DB 2; Length 1500;						
Best Local Similarity 94.6%; Pred. No. 1.5e-223;						
Matches 1319; Conservative 0; Mismatches 75; Indels 1; Gaps						
Qy	1	ggtgtcaaaaaataccaatacttcgtttggcggaactgtgtgccccttgcgtggcaggc	60			
Db	1	gtgttcaaaaaataccaatacttcgtttggcggaactgtgtgccccttgcgtggcaggc	60			
Qy	61	tgcgaaggcagcgagcttttcgtgtcggaacaaaagaagcattccttcgtagacgc	120			
Db	61	tgcgaacaaaggcagcgagcttttcgtgtcggaacaaaagaagcattccttcgtagacgc	120			
Qy	121	atcgacacacccaagacgacgagtcgtcagtatgctgtgcgcgaacttgcgcgaactg	180			
Db	121	atcgacacacccaagacgacgagtcgtcagtatgctgtgcgcgaacttgcgcgaactg	180			
Qy	181	gttcaagcgaaggcccgagcagtcgtcagtatgctgtgcgcgaacttgcgcgaactg	240			
Db	181	gttcaagcgaaggcccgagcagtcgtcagtatgctgtgcgcgaacttgcgcgaactg	240			
Qy	241	aacgcagcgcaatgcgcaaaacattccgaccgcttgcgacagcagcaccgttctac	300			
Db	241	aacgcagcgcaatgcgcaaaacattccgaccgcttgcgacagcagcaccgttctac	300			

[illegible]

GPLNLQGEVVGVINSQIFTRSGGPMGLSFAIPIDVALNVADOLKKAGKVSRCGLGWVI
VGNKDLAESGLDKPSGALVAQLVEDGPAARGLQGVGDVILSLNGQSNESADLPHL
VGNKDPDKINLDIRNGQRKSLSMAGVSLPDDDEEIASMGAPGAERSNRUGVTVD
LTAPQRKSLDLOGGVVKEVQDGPAAVIGLRPGDVIHLDNKAVTSTKVFADVAKALP
KNRSVSMRVLQGRASFTFKLAE"

BASE COUNT 332 a 614 c 657 g 358 t
ORIGIN

Query Match 26.2%; Score 365.6; DB 2; Length 1961;
Best Local Similarity 68.0%; Pred. No. 5.1e-58;
Matches 525; Conservative 0; Mismatches 244; Indels 3; Gaps 1;

Qy	374	tcgggttcgggttcctacatcagcaaaagcgtcatctcgtacccaataccacacgttgctg	433
Db	785	TGGGTTCGGGTTCATCATCTCCACGACGGCTACATCTCCACCAACATCAGCTCGTGG	844
Qy	434	ccggtataggcagtatcaaaagtcctgctcaacgacgaagcgcgaataataccgccaactca	493
Db	845	CCGATGCCGACGAGATCTGTGCGCTGTCCGACCGTACGACGACACAGGCCAAGCTTA	904
Qy	494	tcgggttcggatgtccaatcccgatgtcgccctctgaaatcgacgcaacggaagagctac	553
Db	905	TCGGCGCGACCCGCGACGAGCTGGCGGTGCTGAAGATCCGAGG---CGAAGAACCTGC	961
Qy	554	ccgtcgtcaaaatcggcaatcccaaaaatttgaaacccggcggaatgggtcgctgccaatg	613
Db	962	CGACCTGAACTGGCGCATTCGAAACAGCTGAAAGTGGCGCAATGGTCTCTGGCCATCG	1021
Qy	614	gcgcgaccttcggtttgacaacacgctgacgcgcggcatcgtgtccgcacaagcgagaa	673
Db	1022	GTTCCGCCCTGGCGTTCGATCATCTCGGTACCGCCGCGTATCTGTCAGTGGAGGGTCGA	1081
Qy	674	gcctgcacaaagaaagctacacaccccttcacccaacacgacgttgccatcaatcccgcca	733
Db	1082	GCCTGCCGACGAGAGCTACGTACCTTCATCCACACCGACGTGGCGATCAACCCGGTA	1141
Qy	734	attcgcggcgccgctgttcaacttaaaaggacaggtcgtcggcatcaattccgcaaatat	793
Db	1142	ACTCGCGCGTCCGCTGCTGAACCTGCAGGGCGAAGTGTGCGCATCAACTCGCAGATCT	1201
Qy	794	acagccgacgagcgagatcattgagcatctcctttgccaatcccgattgacgttgccatga	853
Db	1202	TCACCCGTTCCGGCGGCTTCATGGGCGCTGTCTCTCGCCATTCGATTCGATGTCGCGCTA	1261
Qy	854	atgtcgcgacagctgaaacacacgcgcaaaagtccaaacgcggacaaactggcgctgatca	913
Db	1262	ACGTGCGCGACCATTTGAAGAAGCGCGAAGGTGAGTTCGCGGCTGGCTGGCGGTGTA	1321
Qy	914	ttcaggaagtatcctacagtttggcacagctgctcgtgtgtgataaaagccagcgccgcat	973
Db	1322	TCCAGAAAGTGAACAAGGATCTCGCCGAGTCTCTCGGCTCGACAAAGCGCTCGCGCGCG	1381
Qy	974	tgattgccaaaatcctcccgccagcccgacagacgtgcccgcctgcagcgccgcagca	1033
Db	1382	TGGTGGCCCACTGGTGAAGACGCGCGCGCGCGCAAGGGTGGCGCTGCGAGTGGCGCATG	1441
Qy	1034	tcgtcctcagcctgcagcgcgagaaataacgttcttcgcggacaccttcccgctcaggtcg	1093
Db	1442	TGATCTCAGCCTGNACGGCCAGTCGATCAACGAGTCCGCGACCTGCGCGACCTGGTGG	1501
Qy	1094	gcgcattaccgcgggaaaaaagtcagccttcgcgtatgcccgaagaaagcga	1145
Db	1502	GCAACATGAAGCGCGCGCACAAAGATCAACCTGGAGCTGATTCGCAACGCGCCA	1553

RESULT 8
PAU49151 PAU49151 4587 bp DNA BCT 13-MAR-1996
LOCUS Pseudomonas aeruginosa alternate sigma factor (algU), muca, mucB,
DEFINITION mucC and mucD genes, complete cds.
ACCESSION U49151 U24569 U08380 L04794 L02119
VERSION U49151.1 GI:1220191

KEYWORDS
SOURCE
ORGANISM

Pseudomonas aeruginosa strain-PA01.
Pseudomonas aeruginosa
Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
Pseudomonas

REFERENCE
AUTHORS
TITLE

1 (bases 409 to 1055)
Martin,D.W., Holloway,B.W. and Deretic,V.
Characterization of a locus determining the mucoid status of
Pseudomonas aeruginosa: AlgU shows sequence similarities with a
Bacillus sigma factor
J. Bacteriol. 175 (4), 1153-1164 (1993)
93163045
2 (bases 528 to 2836)
Martin,D.W., Schurr,M.J., Mudd,M.H. and Deretic,V.
Differentiation of Pseudomonas aeruginosa into the
alginate-producing form: inactivation of mucB causes conversion to
mucoidy
Mol. Microbiol. 9 (3), 497-506 (1993)
94018645
3 (sites)
Martin,D.W., Schurr,M.J., Mudd,M.H., Govan,J.R., Holloway,B.W. and
Deretic,V.
Mechanism of conversion to mucoidy in Pseudomonas aeruginosa
infecting cystic fibrosis patients
Proc. Natl. Acad. Sci. U.S.A. 90 (18), 8377-8381 (1993)
93391358
4 (bases 252 to 488)
Martin,D.W., Schurr,M.J., Yu,H. and Deretic,V.
Analysis of promoters controlled by the putative sigma factor AlgU
regulating conversion to mucoidy in Pseudomonas aeruginosa:
relationship to sigma E and stress response
J. Bacteriol. 176 (21), 6688-6696 (1994)
95050230
5 (sites)
Yu,H., Schurr,M.J. and Deretic,V.
Functional equivalence of Escherichia coli sigma E and Pseudomonas
aeruginosa AlgU: E. coli rpoE restores mucoidy and reduces
sensitivity to reactive oxygen intermediates in algU mutants of P.
aeruginosa
J. Bacteriol. 177 (11), 3259-3268 (1995)
95286510
6 (bases 1 to 252)
Schurr,M.J., Yu,H., Boucher,J.C., Hilbler,N.S. and Deretic,V.
Multiple promoters and induction by heat shock of the gene encoding
the alternative sigma factor AlgU (sigma E) which controls mucoidy
in cystic fibrosis isolates of Pseudomonas aeruginosa
J. Bacteriol. 177 (19), 5670-5679 (1995)
96032406
7 (bases 1 to 4587)
Boucher,J.C., Martinez-Salazar,J., Schurr,M.J., Mudd,M.H., Yu,H.
and Deretic,V.
Two distinct loci affecting conversion to mucoidy in Pseudomonas
aeruginosa in cystic fibrosis encode homologs of the serine
protease HtrA
J. Bacteriol. 178 (2), 511-523 (1996)
96134987
8 (bases 1 to 4587)
Boucher,J.C., Martin,D.W., Schurr,M.J., Deretic,V., Yu,H., Mudd,M.
and Martinez-Salazar,J.
Direct Submission
Submitted (09-FEB-1996) John C. Boucher, Microbiology, U.T. IISC San
Antonio, 7703 Floyd Curl Dr., San Antonio, Tx 78284, USA
On Mar 11, 1996 this sequence version replaced gi:1173501.
Location/Qualifiers
1..4587
/organism="Pseudomonas aeruginosa"
/strain="PA01"
/db_xref="taxon:287"
/map="67.5"
22
/note="transcription start site for promoter p5; 3' end is
unknown"
124

SOURCE

mRNA

mRNA

```

/note="transcription start site for promoter P4; 3' end is
unknown"
promoter
204..238
/note="P3 promoter"
-35_signal
204..209
/note="P3 promoter"
-10_signal
226..230
/note="P3 promoter"
235
/note="Algu-dependent transcription start site for
promoter P3; 3' end is unknown"
238
/note="Algu-dependent transcription start site for
promoter P3; 3' end is unknown"
310
/note="transcription start site for promoter P2; 3' end is
unknown"
389..394
/note="P1 promoter"
389..421
/note="P1 promoter, the first of two Algu-dependent
promoters; Algu-dependent promoters strongly resemble the
sigmaE promoters of E. coli in both the -10 and -35
sequences"
411..415
/note="P1 promoter"
421
/note="transcription start site for promoter P1; 3' end is
unknown"
474..1055
/gene="algu"
CDS
474..1055
/gene="algu"
/codon_start=1
/transl_table=11
/function="conversion to mucoidy; required for algD
transcription"
/product="sigma factor"
/protein_id="AAC43714.1"
/db_xref="GI:1173502"
/translation="MULTIQDQQLVERVQGRKRAFDLLVLKYHKILGLIVFVHDA
QEAQVQAQAFKAYRALGNFRGDSAFYTWLYRIANTAKNHLVANGRRPPDSQVAE
DAEFFEGDHAKLDIESPERAMLRDEIEATVHOTIQQLPEDLRALTALREPEGLSYEDI
ATVMQCPVGTVRSRIFRAREIDKALQLIREA"
1087..1671
/gene="mucA"
1087..1671
/gene="mucA"
/codon_start=1
/transl_table=11
/product="MucA"
/protein_id="AAC43715.1"
/db_xref="GI:1173503"
/translation="MSREALQETLSAVMNEADELELRRVLAACGDAAELRSTWSRQ
LARSVNHREPTLPKLDIAAASALADEAPPAEKGPWRMVGRLAAVSYTLAVLAG
VRLYNHQPALPQMAQOQTTPQIALPVQKPAVLGAVSEEGAPQVITNSSSDTRWHE
QRPLPYLRHVQQSASVSGTESALPYARAASLENR"
1680..2630
/gene="mucB"
1680..2630
/gene="mucB"
/codon_start=1
/transl_table=11
/product="MucB"
/protein_id="AAC43716.1"
/db_xref="GI:1220192"
/translation="MRTTSLILLGLSMVPAQDAASDWNLRLEADRONSPQGTIF
VYRNGSFSTHEIWRVSDGVRERLIQDGAQEVVRVDGRTQCISGGIADQLADA
QLMPVRFDPQSALSDYDLRLVGSVRAGRPVAVLAVTPRQDHYRGFELHLDRTGLP
LKSLLNKGQLLERFQTLNTGAAPAEQLOQAECQVVGPAKADGEKTVAWRSEW
LPQCFTLTRSMRSPVTPDPVACLTGYDGLARFVFIPLHIGAMVGDARSQLGPTTVV
VSRKLQIDDDGGQVTVVGVVEVPLGTAEARVALSIRPEAAQK"
2627..3082

```

```

/gene="mucC"
2627..3082
/gene="mucC"
/codon_start=1
/transl_table=11
/product="MucC"
/protein_id="AAC43717.1"
/db_xref="GI:1220193"
/translation="MIEQGRVVATEPGAVVETVRRSTCSSANAGCGQGLMQRLG
VCAGRARVPALSDLSIRLVGDAVLGTHEDLLRASVLYFLPLGLGFVAALLATRAGL
VEPLIVSGLAGLLAANLLVRRHARRHADDPASQPVVLRALLSGPDSDA"
3122..4546
/gene="mucD"
3122..4546
/gene="mucD"
/codon_start=1
/transl_table=11
/product="MucD"
/protein_id="AAC43718.1"
/db_xref="GI:1220194"
/translation="MHTLKRKMAAMVALLALSMTARAEPLDFTPLVEQASPAVUNI
STROKLPDRAMARGQLSDPEGLPMFRDFLERSIPOVPRNPRGOOREAQSLGSGFI
ISNDGVILTNHVADADELIVLSDRSEHKAKLIGADPSDVALEAKNLPTLKL
GDSNKLKVGENVLAIGSPFGFHSVAGISAKGRSLPNEYSYVFFIOTDVAINPNSG
GPLLNQGEVYVINSQIFTRSGGFMLGSAIPIDVALNVADQLKAKGVSRGWLGVVI
QEVNKDLAESFGLDKPSGALVAQLVEDGPAAGKGLQVGDVILSINGQINESADLPHL
VGNMKPGDKINLDVIRNGQRKSLSMVAGSLPDDDEETIASMGAFGAERSINRLGTVIAD
LTAQQRKSLDIOGGVVYKEVQDGPAAVIGLRPGDVITHLDNKAVTSTKVPADVAKALP
KNRSVSMRVLROGRASFITFKLAE"
BASE COUNT      811 a 1447 c 1515 g      814 t
ORIGIN

Query Match      26.2%; Score 365.6; DB 2; Length 4587;
Best Local Similarity 68.0%; Pred. No. 4.3e-58;
Matches 525; Conservative 0; Mismatches 244; Indels 3; Gaps 1;

Qy 374 tcggttcgggttcacatcagcaaaacggctacatctcaccataaccacacgtcggtg 433
Db 3411 TGGGTTCCGGTTCATCTCCAACGACGGCTACCTCCACCAACATCAGCTCGTGG 3470

Qy 434 ccggtatggcagatcaaaagtctcgtctcaacgacagcgcaatataccgcgaatca 493
Db 3471 CGGATCGCGAGAGATCTGTCGCGCTGTCGACCCGTAGGAGGACCAAGCCAGTTGA 3530

Qy 494 tcggttcgggttcacatcgcgccttcgtgcaaatcgacgcaacggaagagctac 553
Db 3531 TCGGGCGGACCCGCGCAGGACGCTGGCGGTGCTGAAGATCGAGG---CGAAGAACCTGC 3587

Qy 554 ccgtcgtcaaaatcggcaatccccaaaattgaaccggggcgaaatgggtcgccatcg 613
Db 3588 CGACCTTGAAACTGGCGGATTCGAACAAGCTGAAAGTGGGGGAATGGGTCTCGCCATCG 3647

Qy 614 qcgcgcccttcggctttgacacacagcgtgacgcgcggcatcgtgtcccgaagagcaga 673
Db 3648 GTTCGCGCTTCGGCTTCACCTCCTCGGTCACCGCGGTATCGTCAGTCCCAAGGTCGTA 3707

Qy 674 gcttgcacacgaagctacacaccccttcacacacacgacgcttgcacacacccgggca 733
Db 3708 GCCTGCCGAACGAGAGCTACGTACCTTCATCCAGCAGCAGCTGCGCATCAACCCGGGTA 3767

Qy 734 attccggcgcccgctgttcaacttaaaaggacaggtcgctcgcatcattcgcaaatat 793
Db 3768 ACTCCGGCGGTCGCTGCTGAACCTGCAGGGGGAAGTGTGCGCATCAACTCGCAGATCT 3827

Qy 794 acagcgcgagcgcggtattcattgggcatctcctttgcatcccgattgacgttgcata 853
Db 3828 TCACCCGTTCCGGCGGCTTCATGGGCCCTGCTCCTTCGCCATTCGGATCGCGCTGA 3887

Qy 854 atgtcgccgaacagctgaaaaacacacgggcaaaagtccaaacgacggaacactggcg 913
Db 3888 AGGTCCGCCACCAAGTTGAAGAAAGCGGCAAGCTCAGTGTGCGGCTGGCTGGCGTGTGA 3947

```

gene

[illegible]

TITLE Characterization of the genes coding for the putative sigma factor
Algu and its regulators MucA, MucB, MucC, and MucD in *Azotobacter*
vinelandii and evaluation of their roles in alginate biosynthesis
J. Bacteriol. 178 (7), 1800-1808 (1996)
96178940
2 (bases 1 to 2027)
Martinez-Salazar,J.M., Moreno,S., Najera,R., Boucher,C., Espin,G.,
Soberon-Chavez,G. and Deretic,V.
Direct Submission
Submitted (30-JUN-1995) Jaime M. Martinez-Salazar, Microbiology,
UTHSCSA, 7703 Floyd Curl Drive, San Antonio, TX 78284, USA
FEATURES
Source Location/Qualifiers
1..2027
/organism="Azotobacter vinelandii"
/strain="UW 136"
/db_xref="taxon:354"
1..471
/gene="mucC"
1..471
/gene="mucC"
1..471
/note="similar to Photobacterium sp. ORF4 product, encoded
by GenBank Accession Number L41667"
/codon_start=1
/transl_table=11
/product="MucC"
/protein_id="AAB01512.1"
/db_xref="GI:1345103"
/translation="MIEPGRVVAIERAVWIEALDEACPGCTASVGGCRWSPDWLE
TPRRRRIRALSVMLPRVGDVAVGIRDEFRLSTRALAIYLLPLGLPVCTIAIQOLGF
GESLMIPAGFLGVSVWAMVRISRQMDPTLPQVLRALPQASEQVQRAAF"
483..1904
/note="similar to Brucella abortus HtrA protein encoded by
GenBank Accession Number U07352"
/codon_start=1
/transl_table=11
/function="serine protease"
/product="MucD"
/protein_id="AAB01513.1"
/db_xref="GI:1345104"
/translation="MSKVGFKFSYSLAGALLGSLFVQALPEFTSLVEEASPAVV
NISTRQLPSTVGLPDLGLPLFLREFLESIPLQPRTPDNGRREAHSLGSGFI
ISPDGYLTNNHVVADADEIIVLSDRSELEALVGADPLTRCSFVGGQGFESPHSOT
GTYRPIESRGMCSGHREPERFRSFRDCGHQCHGAKPAERELVPFTOTDVALIPNCSG
GPLDLDRVIGINSQIETRSRGFMGLSFAIEPVMGVADQLKATKGVKARGVLVII
VEGNKDLAESGLDRPAGALVAQVLEDGPDAGGLQVGDVILSDGHPIMVMSADLPHL
QVGLKPGAAANLEVRDGRNRNIAITVGALEPEGNGVQPSIAGTEQSSNRLGVTVTEL
TAEQKSLDLKGGVVIREVLNGPAALIGLRPGDVVTHLNNQPIDSAKTFEAVAGALPK
GRSVSMRVLRQGRASFTFKLAE"

gene	source
CDS	
CDS	
BASE COUNT	366 a 529 c 658 g 474 t
ORIGIN	
Query Match	19.7%; Score 274.8; DB 2; Length 2027;
Best Local Similarity	58.9%; Pred. No. 2.9e-41;
Matches	527; Conservative 0; Mismatches 362; Indels 5; Gaps 3;
Qy	295 tctacgaattttcaaacgctctgctcccgaaacatcccgaaatcccccaagaagaagca 354
Db	
Qy	693 TTCAGGGAGTTCCTGGAGCGCAGCATTCGCGCACTTCCGCGTACTCCGGATAACGGCGG 752
Db	
Qy	355 gatgacgaggattgaacttcggttcggttcacatcagcaaaaaacggtacatctg 414
Db	
Qy	415 accaataccacgctgcttgcggatgggagcagtatcaaatcctgctcaacgacaagcgc 474
Db	
Qy	813 ACCAACACCATGTGTGGCCGATGCCGATGAATCATCTCGTGGCTTGTCCGATCGCAGT 872
Db	
Qy	475 gaataaccgcaaacatcattcggttgatg-tccaatccgattgctgcctcttgaataa 533
Db	
Qy	873 GAGCTCAGGCGCAGTGTGTCGGGCGCATCTCTTACCCGATGTAGCTTTGTTGAAGGT 932
Db	
Qy	534 cgacgcaaggaagcactaccgctgctcaaatccgcaatcccccaaatattgaaacggg 593
Db	

Db	933 CAAGGGTTCGAAT---CTCCCCACAGTCAAACTGGCAGCTACCGACCAATTCAGATCGG 989
Qy	594 cgaatgggtcgctgcctcatcgccgagcgcctctcgccctllgacaacagcgtgacacgcat 653
Db	
Qy	990 GGAATGGGTTCGGCCATCGGTTCCTCCCTTCCTCGATTCATTCGCTGCTGCGGCGAT 1049
Db	
Qy	654 cgtgtccgcaaacgagcagcctgcccacgaaagctacacaccccttcatcccaacgca 713
Db	
Qy	1050 CATCAGTGCACGGGCGAGCCTGCCGACGAGAGTTA-GTTCTTTTCATCCAGACCGA 1108
Db	
Qy	714 cgtgtccatcaatccggggcaattccggcgcccgctgttcaacttaaaaggacggtcgt 773
Db	
Qy	1109 TGTGGCCATCAATCCCGGTAACTCGGCGGCGCTCTTCGATCTCGATCGACGGGTCA 1168
Db	
Qy	774 cggcatcaattcgcaaatatacagccagcagcgaggtattcaggcatctcccttaccat 833
Db	
Qy	1169 AGCATCAACTCCAGATATTCACCCCGTTCGGGGGCTTCATTCGGCTTCGTCATTC 1228
Db	
Qy	834 cccgattgagcttgccatgaatgctgcgcaacagctgaaacacacccgcaaacgcaacg 893
Db	
Qy	1229 TCCCATCGAGGTTCCTATGGCGTGGCGATCAGTTGAAGGCCACTGGCAAGGTTCGTCG 1288
Db	
Qy	894 cggacaactggcgctgattattcagggaagtattcctacggttttggcacagctctcggtct 953
Db	
Qy	1289 CGGTTGGTTGGGAGTAATTCATTCAGGAAGTCAACAAGGATCTGGCTGAGTCTCTTGGTCT 1348
Db	
Qy	954 ggataaagccagcggcgcatgattgccaaatccctcccgacaccccgcaaacgtgc 1013
Db	
Qy	1349 GGATCGCCCGCGGACCGTGTGGTCCGCCAGCTCTTCGAGGATGGACCGCGGACAAGG 1408
Db	
Qy	1014 cggcctgagggcgagcagctgctccagcctgacgagcgaggaataacgtcttcccg 1073
Db	
Qy	1409 CGGTTTGCAGTTCGGCGATGTCATTCACGCTAGATGTCATCTTATGTCATGTCGCG 1468
Db	
Qy	1074 cgacctcccgctcagctgctggcgccattaccccggggaaagcagcctcggcgatcg 1133
Db	
Qy	1469 CGATCTCCCGCATCTGCTTGGGGGGCTCAAAACCCGGGCTCGAGCAATCTCCAGTGGT 1528
Db	
Qy	1134 ggcgaagcgaggaataatcaaatcaaacgcaagctgggcaacgcgcgcgagca 1187
Db	
Qy	1529 CGGTGACGCAAGCGGAGGAACATCGCTATCATCTGTCGGGGCTTCGCGGAGCA 1582
Db	
RESULT 11	
AE004037	12245 bp DNA BCT 06-JUL-2000
LOCUS	Xylella fastidiosa, section 183 of 229 of the complete genome.
DEFINITION	AE004037 AE003849
ACCESSION	AE004037.1 GI:9107394
VERSION	AE004037.1
KEYWORDS	Xylella fastidiosa.
SOURCE	Xylella fastidiosa
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group; Xylella.
REFERENCE	1 (bases 1 to 12245)
AUTHORS	Silvestri,M.L., Siqueira,W.J., de Souza,A.A., de Souza,A.P., Terenzi,M.F., Trufi,D., Tsai,S.M., Tshako,M.H., Vallada,H., Van Sluys,M.A., Verjovski-Almeida,S., Vettore,A.L., Zago,M.A., Zatz,M., Meidanis,J. and Setubal,J.C.
TITLE	The genome sequence of the plant pathogen Xylella fastidiosa. The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing and Analysis, Sao Paulo, Brazil
JOURNAL	Nature 406 (6792), 151-157 (2000)
MEDLINE	20365717
REFERENCE	2 (bases 1 to 12245)
AUTHORS	Simpson,A.J.G., Reinach,F.C., Arruda,P., Abreu,F.A., Acencio,M., Alvares,R., Alves,L.M.C., Araya,J.E., Bata,G.S., Haptista,C.S., Barros,M.H., Bonaccorsi,E., Bordin,S., Bove,J.M., Briones,M.R.S., Bueno,M.R.P., Camargo,A.A., Camargo,L.E.A., Carraro,D.M., Carrer,H., Colauto,N.B., Colombo,C., Costa,F.F., Costa,M.C.R., Costa-Neto,C.M., Coutinho,L.L., Cristofani,M., Dias-Neto,E., Docena,C., El-Dorri,H., Pacinca,A.P., Ferreira,A.J.S.,

TITLE JOURNAL	Ferreira,V.C.A., Ferro,J.A., Fraga,J.S., Franca,S.C., Franco,M.C., Frohme,M., Furlan,L.R., Garnier,M., Goldman,G.H., Goldman,M.H.S., Gomes,S.L., Gruber,A., Ho,P.L., Hoheisel,J.D., Junqueira,M.L., Kemper,E.L., Kitaajima,J.P., Krieger,J.E., Kuramae,E.E., Laigret,F., Lambais,M.R., Leite,L.C.C., Lemos,E.G.M., Lemos,M.V.F., Lopes,S.A., Lopes,C.R., Machado,J.A., Machado,M.A., Madeira,A.M.B.N., Madeira,H.M.F., Marino,C.L., Marques,M.V., Martins,E.A.L., Martins,E.M.F., Matsukuma,A.Y., Menck,C.F.M., Miracca,E.C., Miyaki,C.Y., Monteiro-Vitorello,C.B., Moon,D.H., Nagai,M.A., Nascimento,A.L.T.O., Netto,L.E.S., Nhani Jr.,A., Nobrega,F.G., Nunes,L.R., Oliveira,M., de Oliveira,M.C., de Oliveira,R.C., Palmieri,D.A., Paris,A., Peixoto,B.R., Pereira,G.A.G., Pereira Jr.,H.A., Pesquero,J.B., Quaggio,R.B., Roberto,P.G., Rodrigues,V., de M. Rosa,A.J., de Rosa Jr.,V.E., de Sa,R.G., Santelli,R.V., Sawasaki,H.E., da Silva,A.C.R., da Silva,F.R., da Silva,A.M., Silva Jr.,W.A., da Silveira,J.F., Silvestri,M.L.Z., Siqueira,W.J., de Souza,A.A., de Souza,A.P., Terenzi,M.F., Truffi,D., Tsai,S.M., Tshako,M.H., Vallada,H., Van Sluys,M.A., Verjovski-Almeida,S., Vettore,A.L., Zago,M.A., Zatz,M., Meidanis,J. and Setubal,J.C.		
	Direct Submission		
	Submitted (02-JUN-2000) Organization for Nucleotide Sequencing and Analysis, Bioinformatics Lab - IC/Unicamp, C.P. 6176, Campinas, SP 13083-970, Brazil		
REMARK FEATURES	Location/Qualifiers		
	source	1..12245	
		/organism="Xylella fastidiosa"	
gene		/db_xref="taxon:2371"	
		80..1624	
		/gene="XF2241"	
CDS		80..1624	
		/gene="XF2241"	
		/note="similar to GI 1184684 (percent identity: 47 %/query alignment coverage: 91.2 %/subject alignment coverage: 98.9 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder/Start codon shift: -48"	
gene		/codon_start=1	
		/transl_table=11	
		/product="periplasmic protease"	
CDS		/protein_id="AAF85040.1"	
		/db_xref="GI:9107395"	
		/translation="MNSRIHTRCFGLLAITLPLAACGAOHNSPTPTAPSTPIVTPSTTPAQVLVAGLPDFTQLVDQVPGVGNVETVTRKKVGRKRPIDNDIPEFRFRFGP DFNPNRGQDDGGIAGMGSGFIISKDGILTNHVTIGASEVTIKLTDREFF KAIIGSDQYVALLIDAKNPTVRIGDSSSLKSGQWVAIGSPGLDHSVTAGIV SALSSTSDQRYVFFIQTDVPIINQSGPGLNTRGEVIGINQSFASGQYMGISF AIPNLINAAEQIRKTKGVORSMLGEIGIDALKAQGLGPSRGALVNNIPHSP AAKAGIEVGQVIRSVNGKVLSFSDDLPLIGMPPTGKATGLIIRDKGPRVTVTLTA LNODATASEDDDTAPSKPETSANVELLIGQVENLSAERERLASSQNAKGQVRITAV TAPWANPPLVEGLVILIRIGRTPVSNVAELNRVLNKRKGDVIMLLVTNGRGTFYS ALKAGG"	
gene		1662..1826	
		/gene="XF2242"	
		1662..1826	
CDS		/gene="XF2242"	
		/note="hypothetical protein; identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder"	
		/codon_start=1	
gene		/transl_table=11	
		/product="hypothetical protein"	
		/protein_id="AAF85041.1"	
CDS		/db_xref="GI:9107396"	
		/translation="MRLGVPSFLVNGNTPRPNINIQVFLSTLTILNHAIIILRYPDGGS GAPSPCHPIR"	
		1808..3616	
gene		/gene="XF2243"	
		1808..3616	
		/gene="XF2243"	
CDS		/note="similar to SP P07682 (percent identity: 69 %/query alignment coverage: 99.5 %/subject alignment coverage: 100.0 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder"	
		/codon_start=1	

gene		/transl_table=11	
		/product="GTP binding protein"	
		/protein_id="AAF85042.1"	
CDS		/db_xref="GI:9107397"	
		/translation="MSSDPMRNFRNSIIAHVDHGKSTLADRIITQLCGGLEAREMEAO ALLVDVDSOGVEAQSVANCYTAVEQGLEVPYILNKIDLPTADTERAKREIARVIGIDA QSEITPGAKLLVMSTGRSHPDVAVGVTTPKRKTLAKLTAGEVGWVTASIKDVHGAIPVG DLTLTANDPAKPLPGFQEVQPRFAGLFDVDAEDYDLREALEKRLNDALRFEPE NSEAMGFGRFCFLGMLHMEIVQEREREYDLNLTITAPTVIYELVLTGTLVAMDNP AKMPPINOINEIREPTIRSNLTPPEVGAIVITLCEKRGSOIGITVYLGNOVQVAYEL PMAEVLDFDEFDKLSVTRCVASLDYHFLREQEGPFVFDVTLINGDRVDALSVIYHRHQ AERGRLECEKMKDLPRQMFEDVAIOAIGSOISRSSTVRAKRNVLAKCTGGDLSRK KLLLEKQKEGKKMKQIGRVEIPOEAFVLAVLQIDNK"	
		3717..4517	
gene		/gene="XF2244"	
		3717..4517	
		/gene="XF2244"	
CDS		/note="similar to SP P00803 (percent identity: 40 %/query alignment coverage: 113.5 %/subject alignment coverage: 93.2 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder"	
		/codon_start=1	
		/transl_table=11	
gene		/product="signal peptidase I"	
		/protein_id="AAF85043.1"	
		/db_xref="GI:9107398"	
CDS		/translation="MRMKWFEIILVLTGCTIISLDKLFALKRAARAGLLDTEPP IIDYCRSFPPLVLAAILLSRFIAEPYKIPSSMMPNLLIGDFILVNFAYGLRLPITN TKIIPGKPNQGDVAVFKPKPKPDENWIKRVIGLPGDRIGFGHGTLYINGEPVYILK GQYTGKNAGVDPDTLIEELPNHPHTILESIGRGEREGEVVPPQGYFYMGDNRDN SEDSRWIKTHFLPEQNLGRKAFILWLCNWLCKGSFDPSSRIGTVIR"	
		4514..4927	
		/gene="XF2245"	
gene		4514..4927	
		/gene="XF2245"	
		/note="hypothetical protein; identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder"	
CDS		/codon_start=1	
		/transl_table=11	
		/product="hypothetical protein"	
gene		/protein_id="AAF85044.1"	
		/db_xref="GI:9107399"	
		/translation="MNERLVITKENLVKRTQYGITLRSTLIIFTIIFTFVFAIHAFL VYQKYVSLQTKAMKSLKGLNSDLDPSTIEELFKRINIDSFNVKKNVFKFIRDG WKMKINFEILRXKRGNMDLVGTDEQDLTSRDIK"	
CDS		4959..5597	
		/gene="XF2246"	
		4959..5597	
gene		/gene="XF2246"	
		/note="similar to SP P05797 (percent identity: 50 %/query alignment coverage: 98.1 %/subject alignment coverage: 92.0 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder/Start codon shift: -93"	
		/codon_start=1	
CDS		/transl_table=11	
		/product="ribonuclease III"	
		/protein_id="AAF85045.1"	
gene		/db_xref="GI:9107400"	
		/translation="MFTDPSLLIQLAQRHCSAGTPHNERLEFLGDSVNVNLLIAELFOR WPRAGALTRARSELVRETSLASIAITRMQLGQILIGPGLSGVRRDRIILADAVE AVIAIYLDADLATCTRVLPWFETALTALPVGPKDPKTRLEQARQWSLPVYE LIFESGDPTHKHFVRSCTLCELKLRTEGSGSRRRLAEQDAASHAIDQLDSNK"	
		5594..6490	
gene		/gene="XF2247"	
		5594..6490	
		/gene="XF2247"	
CDS		/note="similar to SP P43728 (percent identity: 51 %/query alignment coverage: 97.3 %/subject alignment coverage: 96.0 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder"	
		/codon_start=1	

```

/transl_table=11
/product="GTP binding protein"
/protein_id="AAF85046.1"
/db_xref="GI:9107401"
/transl_table="MTQTVPCGRIAVIGRPNVCKSTLTNALVGVKISIVSNRPQTT
RHRUGLTAFTPEGQVLVDPGLCHREQKHPKMNLMNRTARGSLSDVAAALLVQSTGHW
NEEDTLAYNLNDGTPVIVLVINKIDRDKDSALPLFLTHINENHFTTTHPYSAKRR
KGLTELVSDLLALPEGDPMSEDEITDQSRFLASELVREQVMQRLGELPVTATVEI
IEXTENTFERICALIWRERESQALVIGKGGARLKETGVKARQOMERLFTQKVLFLE
TWBVRKDWMSNEAAALTKFGYE"
6661..7392
/gene="Xf2248"
6661..7392
/gene="Xf2248"
/note="similar to SPIP5027 (percent identity: 35 %/query
alignment coverage: 97.1 %/subject alignment coverage:
97.5 %); identified by Glimmer/RBSfinder"
located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="DNA repair protein RecO"
/protein_id="AAF85047.1"
/db_xref="GI:9107402"
/transl_table="MLIEHEVAFVLHVPRWRETSLLVEVLTQAYRGLGIARQVGLK
KQTLAAQLQPLQWTFISQIRGELGQLQAEALDTPRLKGEAMLASFYINELLIRLP
PRHAPVNELYLAYSQTRERLRTSDLSLWSRLFRILETILGVGFNLECDANGTPIDLP
ATRYVLDLPEGRRLPSSENNNAKRRDQTGHVLLAARNOIPNTNIDLGLRRSMRAVLI

```

Query Match	18.7%	Score	260.4	DB 1	Length	12425	
Best Local Similarity	56.3%	Pred. No.	8.9e-39				
Matches	573	Conservative	0	Mismatches	431	Indels	
	14	Gaps	4				
QY	159	gctgcccgaacttgc	caactggtc	aaagcgaagcccg	cagtcg	tcaatt---c	214
Db							
QY	238	gTtGCCAGATTT	CACCAACTG	TCGACCAAGT	AGGCGCTGG	CTGTCAACATTGA	297
Db							
QY	215	agcgagccccgcgc	ccacccaacacg	gcagcgcaat	gcgaaacccg	atccgacc	274
Db							
QY	298	CGTCATTAC	CGGTAAAGT	TCGGAAACGT	TCGGCGCAT	ACCGCTTGACAACG	357
Db							
QY	275	cgcttgcgcagcga	cccgcttctaca	gaatttttcaag	agcctcgt	cccgaaacgccc	334
Db							
QY	358	CGAGTTTTTC	CGGCGTTCT	TTTGGCCAGAT	TTCCAGATCG	CGAACCAACCTCG	417
Db							
QY	335	aaatcccccaaga	aagcgcagat	gacgcggat	tgaaactc	ggttcggggttcac	394
Db							
QY	418	---CCAGGAT	GACGAGGAG	TATCGT	GTGTCGGGA	TGGGCTCTGGCTTCAT	473
Db							
QY	395	gcaaaaacgctaca	ctctgaccaaat	ccccacgct	tggttcggg	tatgggcagttca	454
Db							
QY	474	CCAAGACGGCT	ACATACT	CACCAACCAT	CAGTTCAT	CTGTGGCGCAG	533
Db							
QY	455	tcctgctcaacgc	aaagcgcgaa	tataccgc	caaatcatc	ggttcggatgc	514
Db							
QY	534	TCAAAATTGACT	GC	CGCGGAAT	TCAAGCAAAAT	TCATCGGTAG	593
Db							
QY	515	atgtgcgcctctg	aaataatcg	acgcgaacg	agctaccc	gctgcataatccg	574
Db							
QY	594	ATGTGCTTTACT	CAAAATCG	ATCGACAA	---AAAATC	TACCCACCGTCG	650
Db							
QY	575	ccaaaatttgaac	accgggcgaa	tggtgcgt	gccatcg	gcgcgccttcg	634
Db							
QY	651	CCAGTTACAT	TAAATCGGGT	CAGTGGGT	GTGGATCG	CGCTCCCGCTTC	710
Db							
QY	635	acagcgtgacgc	gcggcatc	ggtgtccg	cc---aaagc	agaagcctgcc	691
Db							
QY	711	ACTCAGTCACT	GCAGGGAAT	ATGTCAGCG	CCCTGGGACG	CAGCAGCCAGT	770
Db							
QY	692	acacacccttcac	taaacacgag	tgcattcc	atcggg	caattccg	751
Db							
QY	771	ATGTCCCATTT	ATCCAAACCG	ACGATCCCAT	TACAGGGGA	AAATCTCTGGT	830
Db							
QY	752	tcaacttaaa	agacaggt	tcgtcgcat	caattc	gcgcaatata	811
Db							

[illegible]

Db	523	TTTTTCTCGGCTCGGCCAGACCGTCACTTCCGGCATCGTCACTGGTGCAGCGGCCTC	582
QY	679	CCCAACAAAGCTACACACCCCTTCATCAACACCGAGTTCGATCAATCGGCAATCGGCAATCC	738
Db	583	AACATCGAGAACCTGGAAACATTCATCAGACGATGCGGCGCATCAATCGGCGCAACTCG	642
QY	739	ggcgccgcgtgttcaacttaaaaggacaggtcgctcgccatcaattcgcaatatatacag	798
Db	643	GGTGGGCCCTGCTCAACCTCGCGGCGAGCTGATCGGCATCAACACCGCCATCTCTGGGC	702
QY	799	cgacgcgcgatctaatcgggcatctcttggccatcccgatgaagttgccaatgaatgc	858
Db	703	CCCAACGCGGCAACATCGGATCGCTTGTGCAATCGGCTTCCCAACATGTTGGCGCACCTG	762
QY	859	gccaaacagctgaaaaacaccgcgcaaaagtccaacgcgacgaactggcgatgattattcag	918
Db	763	TCCGAACAGATCGTGAATACGGTGAAGTCGCTCGCGGCGACGCTGGGCATCATAGGGACC	822
QY	919	gaagtactcaggtttggcagctcgcttgcgtctggataaaagcagcgcgcatgatt	978
Db	823	GAGCTCACTCTGAAGTGGCCACAGACCTTCGGCTACACAAAGAGGACGCGGCTTCGTC	882
QY	979	gccaaatctctccgcgagccgcgcaagaagtcgcgccctgcagcgcgagcgacatcgtcgcgcgc	1038
Db	883	AACCAAGGTGATGCTCTGCTCGCGCGCCAAAGCGCGCATCAAGCGCGGCGACATCATC	942
QY	1039	ctcagctcgcagcgcggaataatcgttcttccttcggcgacacttcccgatggtcgcgcc	1098
Db	943	GTCAGCATCGACGCAAGGCCATTCGCTTCCTCGCGAGCTCGCGGCCAAGATTGGCAC	1002
QY	1099	attacgcgggaaagaagtcagctcgcgctgtagcgcaaggcggaagaatcacaatc	1158
Db	1003	ATGGTGCGGCACAGCAGGTGGCATGGCCCTGATCCGCGATGGCAAGGAGGACCGCTC	1062
QY	1159	aaagccaagctgggcaacgccgcoga	1184
Db	1063	AAGGTCACCTTGAAGAGCGGATGA	1088
RESULT	14		
AE004859			
LOCUS	AE004859	11461 bp	DNA
DEFINITION	Pseudomonas aeruginosa PA01, section 420 of 529 of the complete genome.	BCT	30-AUG-2000
ACCESSION	AE004859		
VERSION	AE004859.1	GI:9950678	
KEYWORDS			
SOURCE	Pseudomonas aeruginosa.		
ORGANISM	Pseudomonas aeruginosa		
REFERENCE	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;		
AUTHORS	1 (bases 1 to 11461) Stover, C.K., Pham, X.-Q., Erwin, A.L., Mizoguchi, S.D., Warren, P., Hickey, M.J., Brinkman, F.S., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G.K., Wu, Z., Paulsen, I.T., Reizer, J., Saier, M.H., Hancock, R.E.W., Lory, S. and Olson, M.V.		
TITLE	Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen		
JOURNAL	Nature 406 (6799), 959-964 (2000)		
MEDLINE	20437337		
REFERENCE	2 (bases 1 to 11461)		
AUTHORS	Stover, C.K., Pham, X.-Q., Erwin, A.L., Mizoguchi, S.D., Warren, P., Hickey, M.J., Brinkman, F.S., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R.M., Smith, K.A., Spencer, D., Wong, G.K., Wu, Z., Paulsen, I.T., Reizer, J., Saier, M.H., Hancock, R.E.W., Lory, S. and Olson, M.V.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-MAY-2000) Department of Medicine and Genetics, University of Washington Genome Center, University of Washington		

FEATURES	Box 352145, Seattle, WA 98195, USA	
	Location/Qualifiers	
source	1. .11461	CDS
	/organism="Pseudomonas aeruginosa" /strain="PA01" /db_xref="taxon:287" 203..1306 /gene="mltB1" /note="PA4444" 203..1306 /gene="mltB1" /codon_start=1 /transl_table=11 /product="membrane-bound lytic transglycosylase" /protein_id="AAG07832.1" /db_xref="GI:9950679" /translation="MRRATALPLFLVLSACSSEPTPPKPAAKPQARTVISPVRQ SVQIPILRPGDYANPAQAHPIDRMVSOHGNROQLHDLFAQTORLDMVILMRDQAP TYTPSPNGAMLYRKFFVTPGNQGVLEMDQYEDLORASVYGVPELIVGIIG VETWRGVMGKTRIIDALSTLSFSYPRAEFSFELEOLFLLQARKEGDDPLALGSGY GAMYGQFMPSFTKYAVDFDGDGHIDLWNPDRDAIGSVANYFKQHGWSGGRVAVPAS GRAPSLDEGKFTLYPLDVLASAGLRPQGLGGRHQASLLRLDMDGRNYQYWGLPNFYV ITRYNHSTHYAMAVWEIJCVEKDVRRHRSVVRQD" complement(1293..2051) /gene="PA4445" complement(1293..2051) /gene="PA4445" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAG07833.1" /db_xref="GI:9950680" /translation="MATAISLVLVEADRYLDAARIQDYCPNGLOVEGRPQVRRIVSGV TASQALLDAADVAVLVVHHGVFWKGENPCVVMKQRRRLKTLINNDISLLAYHLPL DLHEVNGVOLARQLEFVGPELPGNPRISVILGSLAEPMQADFAHRYRDALGRE PLLYDAGQTRIAWCTGGAQGYIDQATAAGVDAYLTGEVSEQTVHSARENGISFIAA GHHATERYGVALGDYLGKRAIEHLFIDCPNP" 2168..3337 /gene="algW" /note="PA4446" 2168..3337 /gene="algW" /codon_start=1 /transl_table=11 /product="AlgW protein" /protein_id="AAG07834.1" /db_xref="GI:9950681" /translation="MPKALRPLGHPVLVGLVLLALLIIQHNPVLGLPRQEVHVEQAPL LSRQEGFVSYANAPRAVANLYTKMVKSFHPLFDPMFRFRFGDNLPOQKRM ESSLGSAVIMSAGYLLTNHVTAGADIIVALLRQDRETIAQLVGSQDETDLADID LKNPLAMTGRSDGLRTGDVGLICNPFPGVQOTVMGTISATGRNQLGNTYEFQIT DAALNPGSGALVDAGNLGINTATFSKSGSQGIFATPKLAEVMQSIIEHQO VIRGLGVVKALVPELAESLGLGETAGIVVAGVYRQGPARGGLIPCDVILTIDQOE ASDGRSRNQVARTRPQKISIVVLRNGQKNLTAEVGLRPPPPAPQOQKQDGGE" complement(3380..4435) /gene="hisC1" /note="PA4447" complement(3380..4435) /gene="hisC1" /codon_start=1 /transl_table=11 /product="histidinol-phosphate aminotransferase" /protein_id="AAG07835.1" /db_xref="GI:9950682" /translation="MSKFWSPFVKDLVPYPGEQPKLSRLVKLNTNENPYGPSQALA AMQELMNDLRLYPDNGERLQKVAAHYGVOANQVFGNSDEVLAHFHGLFOHDL PLPEPDVTYSPYVYGLYGAHEKIALDERFRIVERDYARNPGGIIPPNAPFTGCL LPDLAIEMALKASDVSVVYDEAYDFGGEISAIALVDYRPNLLVTQITLSKRSRAGLR VGLAVGHADLVLEALRIKNSFNSPLDRLAJAGAAADEFDAYFRRTCOAVIDSREAL SASIQALGFVELPSAANFVARPHRDAGQIATSLIREOGVIVRHFQKARIQDFLRITI GSPQNALDLALHFLK" complement(4438..5760) /gene="hisD"	
gene	203..1306	gene
	/note="PA4448" complement(4438..5760) /gene="hisD" /codon_start=1 /transl_table=11 /product="histidinol dehydrogenase" /protein_id="AAG07836.1" /db_xref="GI:9950683" /translation="MTAPFAIRRLNAADPDFGRHLDHLLSWESVSDSDSYNQRVLIDIIA AVSRGDAAVVEFTQFDCLOAASMDIILIPRELALRITITVAOREALEVAERVR SYHEKQKCSWRYTEADTVLGOQVTPILDRAGLYVPGGKASYPSSVLMNAIPAKVAGY SEVMNVTPRGEINEIIVLAAACIAGVDVFTTIGCAQVAALAYGTESVPRVDKIVGP GNIYVATAKRHFVGQVGDIMAGPSEILVYCDGQDTPDWIAMDLFSQAEDSDAQISIL VSPDAFLDRVADSIRLPLTPTMERAEIITLSUEBGRGALIQVDAQACAVANRIAPEH LELSVADPESWLPEIRHAGAIFPMGYTAEALGDYCAGPNHVLPTSGTARFSSPLGVYD FOKRSSIIINCSAEGASVLCRTASVLARGESLTAHARSAEYRILDEKEA" complement(5886..6521) /gene="hisG" /note="PA4449" complement(5886..6521) /gene="hisG" /codon_start=1 /transl_table=11 /product="ATP-phosphoribosyltransferase" /protein_id="AAG07837.1" /db_xref="GI:9950684" /translation="MLTIALSKGRILDDTLPLAAAGIVPSNPDKSRKLIPTSLSD VRLLIIVRATDPTVYVEHGAADIGVAGKDVLMYEGGQGLYEPILDLRTANCLMTAGAIG APEPKGRIVATKFEVNVAKRYVAEQGRQVDVILKLYGSMELAPLVGLADKIIDVVDITGN TLRANGLPEQELIATISSRLVYNKASMKMQHGRIOSLIDTLDRDAVEARH" complement(6539..7804) /gene="mura" /note="PA4450" complement(6539..7804) /gene="mura" /codon_start=1 /transl_table=11 /product="UDP-N-acetylglucosamine 1-carboxyvinyltransferase" /protein_id="AAG07838.1" /db_xref="GI:9950685" /translation="MDKLIITGNGRLDGEIRISGAKNSALPILATLLADTPTVTCNL PHLHDITMIELFGRMGVQPIIDEKLNVEVDASSIKTLVAPYELVKTMRASILVLGPM LARFEAEVLPGCCAIGSRPYDLHRLGEANQATIEVEGGYIKAKAPGURGGHFF PDTVSVGTENLMMAAALANGRTVLQNAAREPVDLANCLAMGANVQAGSDTIIV EGVKRLGARGVDVLPDRLETSTYLVAAAATGGRVKLKDTPDTILLEAVLQKLEAGAH STGSNWIELDMKGNBPKANVPTAPYAPPTDMQAOQISMNAAVSGTGAVIETVFNRR FMHVYEMRMNQAILVEGNTAIVTGVPKLGPAPVWATDURASASLVIAGLVAEGDTLLI DRHYHDGREGICEELQLLGGAKIRRVPG" complement(7825..8064) /gene="PA4451" complement(7825..8064) /gene="PA4451" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAG07839.1" /db_xref="GI:9950686" /translation="MQAEDVKSLLQELPCTOVEVEGECNFCOLNLSDELQAGSPVK RQQYAHNLNAWIADGSIHAVTMTKFFSRAAWERS" complement(8181..8489) /gene="PA4452" complement(8181..8489) /gene="PA4452" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAG07840.1" /db_xref="GI:9950687" /translation="MSQASIRGAAGELQALGVLDYSSGPALREOGGRLIRASQAEL VYDCAVERSSSVGISLLAFIRDKARKGVLSVRALPDDMREIAKVSLLLEILPLOQE" complement(8486..9133)	
gene	203..1306	gene
	/note="PA4451" complement(8486..9133) /gene="hisD"	

[illegible]

Db 643 GCCTCCGGCTTCTTCAATCACCGAAGACGGTTACCTCGTCACCAACAACACACGTCGTCTCC 702
Qy 436 ggtatgg-gcagtatcaaatcctgtctcaacgacgaagcgcgaatatataccgcgaactcat 494
Db 703 GACGGCTCGGACCTTACCGTTATCATGAACGACGGAACCGAGCTCGACGCCAAGCTCGT 762
Qy 495 cggttcggatgtccaatccgatgtgcgcctctctgaaaaatcgacgcaaacggaagatcacc 554
Db 763 CGGCAAGGACACAGCCGACCGACTTTCGCGTGTCTAAGGTCTGATGACAAGCGCAAGTTTAC 822
Qy 555 cgtctcaaaatcgcgaatcccaaaattgaaaccggcgaatgggtcgctgcctcgcg 614
Db 823 CTATGTACGTTCCGGGACGACGAAGGTGCGGTGCGCGACTGGGTGGTCTGCTGTCCG 882
Qy 615 cgcgccttcggctttgacaacagcgtgaccgcgcgcctcgtgcgcgaagggcagaag 674
Db 883 CAATCCCTTCGGGCTCGGGCAGCGTCACTGCCGCGATCATCTCCGCTCGGGGCGCGGA 942
Qy 675 cctgcccacgaagactacacacccttcctcaaaacgcagcgttgccatcaatcccgggcaa 734
Db 943 TATCGGCTCCGGTCCCTATGACGATTATCTGAGGTCTGACGCGGTGAACCGCGGCAA 1002
Qy 735 tccggcgcccgctgttcaacttaaaagacaggtcgtcgacatcaattcgcaaatata 794
Db 1003 TTCCGGTGGTCCGACCTTCAATCTTCCGGCGAGGTGGTTCGGTATCAACACGGCCATCTT 1062
Qy 795 cagcgcgagcgggattcatggtgcatctcttgcacatcccgattgacgttgccatgaa 854
Db 1063 TTCGCCCTCGGGCGCAATGTCGGTATCGCTTTGCAATCCCGCGCTCGTTCGGAAGGA 1122
Qy 855 tgcgcggaacagctgaaacacacgcggaagtcgaacgcggaacactggcgctgattat 914
Db 1123 TGTCTGGACTCCCTCATCAAGGACGGCACGGTGTGCGCGGCTGGCTTGGTGTCCAGAT 1182
Qy 915 tcaggaagtatcctacggtttggcacagcttcggtctggtatgataaagccagcgcgcat 974
Db 1183 CCAGCCGTGACGAAGGATATCGCCGAGTCTGCTCGCCCTTTCGGAAGCCGAACGGCCCT 1242
Qy 975 gattgccaaaatccttcccgcgagcccgacgaacgctgcgcgctgcagcgcgcgacat 1034
Db 1243 GGTCTCGAGCCTCAGCGGGGCTCGCGGGTGAAGAGCCGGGATCAAGAACGGCGACGT 1302
Qy 1035 cgtcctcagcctcagcgcgagagaataagttcttccgcgcgaccttcccgtcatggtcgg 1094
Db 1303 GGTGACGGCACTCAACGGGAGCCGCTCAAGGATCCGCGGGACCTCGCACGTTCGCTTC 1362
Qy 1095 cgcattacgcgcgggaaagaagtacgacctcggcgtatggcgcaagcgcaagaatcac 1154
Db 1363 GGCCCTGCGCCCGCTCCACTGCACAGGTTACCCCTGTGGCGCAGCGGCAATCCGAAC 1422
Qy 1155 aatcaaaagccagctgggca---acgcgcgagacataccggcgcatcattccaaaacaga 1211
Db 1423 GGTCAACCTCGAGATCGGCACGCTGCCGAGCTGCAAGGAACCCGACCCGCGACCGG 1482
Qy 1212 tgaagccctctacacgaacagcaatcggtagtctctcgttcgaatccgcaggeattac 1271
Db 1483 TGAAGCGGACGCCGACGAAGGCGAGGCGAGGAGGCGCTGGCCGACCTTCGGCCTGAC 1542
Qy 1272 cct----tcagacacataccgacagcagcgcaaacacctcgtcgttcgtacgggtttccga 1328
Db 1543 GGTGACCCCGTCGGAGGACGGCAAGGGCTCACGATCGCCTCCGTCTGATCCGGATTCCGA 1602
Qy 1329 cgcggcagaacgcgcaggtttaaggcaagcgacgaataatcctagccgtca 1378
Db 1603 GCGCGGATCGC---GGCCTGAAGGAGGCGGAAGATCGTCTCCGTCA 1649

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2001, 03:20:15 ; Search time 91.55 Seconds
(without alignments)
5724.188 Million cell updates/sec

Title: US-09-388-090-3
Perfect score: 1395
Sequence: 1 gtgtcaaaaataccaata.....tcagggaagtcctccgtcaa 1395

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues
Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36.*
1: /cgn2_2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /cgn2_2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /cgn2_2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /cgn2_2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /cgn2_2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /cgn2_2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /cgn2_2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /cgn2_2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /cgn2_2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /cgn2_2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /cgn2_2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /cgn2_2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /cgn2_2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /cgn2_2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /cgn2_2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /cgn2_2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /cgn2_2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /cgn2_2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /cgn2_2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /cgn2_2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /cgn2_2/gcgdata/geneseq/geneseq/NA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1395	100.0	1395	21 293414	NGSP polypeptide c
2	1393.4	99.9	1395	21 251538	Neisseria meningit
3	1373.4	98.5	1500	21 254510	Neisseria gonorrhe
4	1324.4	94.9	1326	21 251539	Neisseria meningit
5	1320.6	94.7	1500	21 233307	Neisseria meningit
6	1315.8	94.3	1500	21 254511	Neisseria meningit
7	1313.8	94.2	1500	21 254512	Neisseria meningit
8	1263	90.5	1500	21 233306	Neisseria meningit
9	1258.2	90.2	1500	21 233305	Neisseria meningit
10	1242	89.0	1242	21 293415	NGSP polypeptide c
11	1176.2	84.3	1347	21 251533	Neisseria meningit
12	1061	76.1	1110	21 233308	Neisseria meningit

13	364.8	26.2	1436	20	X98241	Nucleotide sequenc
14	260.2	18.7	390	21	Z33309	Neisseria meningit
15	176.4	12.6	1854	19	V29528	Escherichia coli s
16	175.6	12.6	1980	12	Q14416	S. typhimurium htr
17	154.2	11.1	1791	17	T04403	Heat shock protein
18	153	11.0	153	21	Z93416	NGSP coding sequen
c	150.4	10.8	1038602	20	Z01425	Complete genome se
20	139.4	10.0	2894	17	T10423	H. influenzae SB33
21	138.6	9.9	153	21	Z51542	Neisseria meningit
c	137.8	9.9	1830121	17	T42063	Haemophilus influe
23	128.8	9.2	1529	19	X14164	H. pylori GHPO 536
24	124	8.9	1540	20	V90845	Nucleotide sequenc
25	120.8	8.7	1484	20	V90921	Nucleotide sequenc
c	110.8	7.9	558	19	V31239	E. coli J96 pathog
27	105	7.5	1185	21	Z29195	cDNA encoding C-te
28	105	7.5	1443	21	Z29176	Recombinant cDNA e
29	105	7.5	1705	20	Z41293	Human normal ovari
30	105	7.5	2036	19	V29540	Homo sapiens D8725
31	105	7.5	2075	18	T75444	cDNA encoding oste
32	105	7.5	2205	19	V04680	Human presenilin 1
33	85	6.1	607	20	X29130	Polynucleotide RTP
34	81.6	5.8	2040	21	Z52362	NGSP gene-14 assoc
35	80.6	5.8	1302	21	Z48700	BAS011 coding seq
36	80.6	5.8	1302	21	Z48701	BAS011 coding seq
37	80.6	5.8	1302	21	Z48702	BAS011 coding seq
c	71	5.1	284	20	V90593	Nucleotide sequenc
39	68	4.9	1553	19	V43618	Human secreted pro
40	62.8	4.5	273	20	V90690	Nucleotide sequenc
41	60.6	4.3	1194	21	Z91868	Streptococcus pneu
42	60.6	4.3	15747	20	X13018	Enterococcus faeca
43	60.6	4.3	21338	19	V52153	Streptococcus pneu
44	60.4	4.3	9848	20	X20503	Polynucleotide seq
45	57.6	4.1	1835	19	V29524	Homo sapiens Psp1

ALIGNMENTS

RESULT	1
293414	
ID	Z93414 standard; DNA; 1395 BP.
XX	Z93414;
AC	Z93414;
DT	24-JUL-2000 (first entry)
XX	NGSP polypeptide coding sequence.
DE	NGSP polypeptide coding sequence.
XX	NGSP; polypeptide; peptide; vaccine; immune response; antibody;
KW	cellular matrix; adherence domain; ligand; detection; diagnosis;
KW	screening; probe; primer; prophylaxis; therapy; ss.
XX	Neisseria gonorrhoeae.
OS	Neisseria gonorrhoeae.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1395
FT	/*tag= a
FT	/product= NGSP polypeptide
XX	
PN	WO200012133-A1.
XX	
PD	09-MAR-2000.
XX	
PF	01-SEP-1999; 99WO-US20070.
XX	
PR	01-SEP-1998; 98US-0098685.
XX	
PA	(ANTE-) ANTEX BIOLOGICS INC.
XX	
PI	Jackson WJ, Harris AM;
XX	
DR	WPI; 2000-237782/20.
DR	P-PSDB; Y83150.

XX Non-cytosolic NSGP polypeptide and polynucleotide sequence from
PT Neisseria useful for diagnosis, prevention or treatment of Neisseria
PT infections
XX
XX
PS Claim 38; Page 60-61; 68pp; English.
XX
CC The NSGP polypeptide of *N. gonorrhoeae* has conserved Arg-Gly-Asp and
CC Arg-Gly-Asn motifs near the C-terminus which function as adherence
CC domains for extracellular matrix proteins. Using the NSGP polypeptide
CC as a vaccine produces antibodies which inhibit binding of *N.*
CC gonorrhoeae to the host's cellular matrix reducing attachment and/or
CC subsequent invasion. The NSGP polypeptide and its peptide fragments
CC can be used to immunise an animal and produce an immune response.
CC They can also be used as ligands to detect antibodies elicited in
CC response to Neisseria infections and also as antigens or immunogens
CC for inducing Neisseria-specific antibodies which are useful in
CC immunosays to detect Neisseria in biological specimens. Nucleotides
CC encoding NSGP or its fragments can be used as probes to identify
CC Neisseria in biological specimens by hybridization or polymerase
CC chain reaction amplification. The NSGP polypeptide can also be used
CC in screening assays to identify agents and compounds which useful as
CC diagnostic, prophylactic or therapeutic agents against Neisseria
CC infection.
XX
SQ Sequence 1395 BP; 359 A; 438 C; 353 G; 245 T; 0 other;

Query Match 100.0%; Score 1395; DB 21; Length 1395;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gtgttcaaaaaataccaataacttcgttggcggcactgtgtgcgccttgcgtggcaggc 60
DB 1 gtgttcaaaaaataccaataacttcgttggcggcactgtgtgcgccttgcgtggcaggc 60

QY 61 tgcgaaaaaggcaggcagcttttcgttcgggacacaaaaaagaagatccttcgtagaacgc 120
DB 61 tgcgaaaaaggcaggcagcttttcgttcgggacacaaaaaagaagatccttcgtagaacgc 120

QY 121 atcgaaacacacaaagacgagcgagtgctcagatgctgtcgtcccgactttgcccacactg 180
DB 121 atcgaaacacacaaagacgagcgagtgctcagatgctgtcgtcccgactttgcccacactg 180

QY 181 gttcacaagcgaagcccgccgagctgcataatttcaggagagcccccgcgcgcaccccaa 240
DB 181 gttcacaagcgaagcccgccgagctgcataatttcaggagagcccccgcgcgcaccccaa 240

QY 241 aacggcagcggaatgcgaaacacgacatccgacccgcttcgcgacagcagccgcttcac 300
DB 241 aacggcagcggaatgcgaaacacgacatccgacccgcttcgcgacagcagccgcttcac 300

QY 301 gaattttcaaacgctcgtcccgaaacatgcgcgaatcccccgaagaagacagatgac 360
DB 301 gaattttcaaacgctcgtcccgaaacatgcgcgaatcccccgaagaagacagatgac 360

QY 361 ggcggattgaacttcgttcgggttcataatcagcaaaaaacggctacatcctgaccaat 420
DB 361 ggcggattgaacttcgttcgggttcataatcagcaaaaaacggctacatcctgaccaat 420

QY 421 accacgctgttcgggtatggcagatcaaaagtccctgctcaacgacagcgcgaatat 480
DB 421 accacgctgttcgggtatggcagatcaaaagtccctgctcaacgacagcgcgaatat 480

QY 481 accgcaaaactcactcgttcggtatccaaatccgctgctgccttcctgaaatcgacgca 540
DB 481 accgcaaaactcactcgttcggtatccaaatccgctgctgccttcctgaaatcgacgca 540

QY 541 acggaagagctaccgctgcataaatccggaatcccaaaaaatttgaaacccggcggaatgg 600
DB 541 acggaagagctaccgctgcataaatccggaatcccaaaaaatttgaaacccggcggaatgg 600

QY 601 gtgcgtgccatgcggcgcccttcggttttgacacagcgtgacgcgcgcatcgtgtcc 660

DB 601 gtgcgtgccatgcggcgcccttcggtttgacacagcgtgacccggtatcgtgtcc 660
QY 661 gccaaaggcagaagccttcgcccacgaaagctacacaccttcacccaacgacgttccc 720
DB 661 gccaaaggcagaagccttcgcccacgaaagctacacaccttcacccaacgacgttccc 720
QY 721 atcaatccgggcaattccggcgcccgctgttcaacttaaaaggagacagctcgtcgcac 780
DB 721 atcaatccgggcaattccggcgcccgctgttcaacttaaaaggagacagctcgtcgcac 780
QY 781 aattcgcaaatatcacgcccgagcggggatttcaggcatctcctttgccaatcccgatt 840
DB 781 aattcgcaaatatcacgcccgagcggggatttcaggcatctcctttgccaatcccgatt 840
QY 841 gacgttccatgaatgctcgcgaaacagctgaaacacccgcaaaagtcacaacgcgacaa 900
DB 841 gacgttccatgaatgctcgcgaaacagctgaaacacccgcaaaagtcacaacgcgacaa 900
QY 901 ctggggcgtgattattcagggaagtacctacggtttggcacagtgctcgggtctgataaa 960
DB 901 ctggggcgtgattattcagggaagtacctacggtttggcacagtgctcgggtctgataaa 960
QY 961 gccagcggcgcatgattgcacaaatccttcggcgagcccccgcagaacgctgcgcgctg 1020
DB 961 gccagcggcgcatgattgcacaaatccttcggcgagcccccgcagaacgctgcgcgctg 1020
QY 1021 caggcggcgacatcgtcctcagcctcagcgcggcgagagaaatacgttcttcggcgacctt 1080
DB 1021 caggcggcgacatcgtcctcagcctcagcgcggcgagagaaatacgttcttcggcgacctt 1080
QY 1081 ccggtcatggtcggcgccattacgcgcgggaaagaaagtcagcctcggtatgcgcgcaa 1140
DB 1081 ccggtcatggtcggcgccattacgcgcgggaaagaaagtcagcctcggtatgcgcgcaa 1140
QY 1141 ggcgaaagaatcacaaatcaaaagccaaagctgggacgcgcgcggagcataccggcgacaa 1200
DB 1141 ggcgaaagaatcacaaatcaaaagccaaagctgggacgcgcgcggagcataccggcgacaa 1200
QY 1201 tccaaaacagatgaagcccccctacacggaacagcaatccggtacgttctcgttcgaatcc 1260
DB 1201 tccaaaacagatgaagcccccctacacggaacagcaatccggtacgttctcgttcgaatcc 1260
QY 1261 gcaggcattacccttcagacacataccgcagcagcgggcaaacacctcgtcgtacagg 1320
DB 1261 gcaggcattacccttcagacacataccgcagcagcgggcaaacacctcgtcgtacagg 1320
QY 1321 gtttcgacgcggcagacgcgcaggttaaggcacgcgcgagaaatcctagccatcagg 1380
DB 1321 gtttcgacgcggcagacgcgcaggttaaggcacgcgcgagaaatcctagccatcagg 1380
QY 1381 gcaagtcgccgtcaaa 1395
DB 1381 gcaagtcgccgtcaaa 1395

RESULT 2
Z51538
ID Z51538 standard; DNA; 1395 BP.
XX
AC Z51538;
XX
DT 03-JUL-2000 (first entry)
XX
DE Neisseria meningitidis NMASP protein-2 encoding DNA.
XX
KW NMASP: non-cytosolic; antibacterial; anti-inflammatory; cytotoxic;
KW anti-NMASP antibody; vaccine; diagnosis; therapy; prophylaxis;
XX Neisseria infection; meningitidis; septicaemia; ds.
OS Neisseria meningitidis.
XX
FH Key Location/Qualifiers

Qy	421	accacgctcgctg	ccggtatg	ggcagtat	caaaagtc	ctgctcaaacgacaaagcgcgaatat	480
Db	421	accacgctcgctt	accgcatg	ggcagta	tcaaaagtc	ctgctcaaacgacaaagcgcgaatat	480
Qy	481	accgccaatac	tcggttc	ggtatg	tccaatcc	cgatgtcgccctctgaaatacgacga	540
Db	481	accgccaatac	tcggttc	ggtatg	tccaatcc	cgatgtcgccctctgaaatacgacga	540
Qy	541	acggaagagct	taccctgt	tcaaaatcg	gcaatcccaaaaa	ttgaaaccggcgcaatgg	600
Db	541	acggaagagct	gcccgt	tcaaaatcg	gcaatcccaaaagt	ttgaaaccggcgcaatgg	600
Qy	601	gtcgctccat	tcgcgccct	tcggcttt	gacaacag	ctgacccgcggcatcggtcc	660
Db	601	gtcgcgcca	tcggcgccct	tcggcttc	gacaacag	ctgacccgcggcatcggtcc	660
Qy	661	gccaaagcaga	gcctg	cccaagaa	gagctacacac	cttcataccaaacgcagcttgc	720
Db	661	gccaaagcaga	gcctg	cccaagaa	gagctacacac	cttcataccaaacgcagcttgc	720
Qy	721	ataatccgg	gcaattcc	ggcgccg	ctgttcaact	taaaaggacaggtctgctggcatc	780
Db	721	ataatccgg	gcaactcc	ggcgccg	ctgttcaact	taaaaggacaggtctgctggcatc	780
Qy	781	aatlcgcaaa	tatacag	cgcgag	cggttc	atgggcatctctttgccatcccgatt	840
Db	781	aactcgcaaa	tatacag	cgcgag	cggttc	atgggcatctcttcgcatcccgatt	840
Qy	841	gacgttgcca	tgaatg	tcgcca	gacagctg	aaacacacggcaaaagtccaa	900
Db	841	gacgttgcca	tgaatg	tcgcca	gacagctg	aaacacacggcaaaagtccaa	900
Qy	901	ctggcgctga	tattcaga	gaagta	tcttcac	ggtttggcacagtctgtcggtctggataaa	960
Db	901	ctggcgctga	tattcaga	gaagta	tcttcac	ggtttggcacagtctgtcggtctggataaa	960
Qy	961	gccagcgcg	catgtat	tgccaaa	atccttc	ccgcgcagcccgcgcagaa	1020
Db	961	gccgcgcg	cactgat	tgccaaa	atccttc	ccgcgcagcccgcgcagaa	1020
Qy	1021	caggcgcg	gacatg	ctctcag	ctcgac	gcgcggaagaatacgttctccggcgacct	1080
Db	1021	caggcgcg	gacatg	ctctcag	ctcgac	gcgcggaagaatacgttctccggcgacct	1080
Qy	1081	cccgctcat	tggtcg	cgccattac	gcgcggg	aaagaagtcagctcggtatggcgcaaa	1140
Db	1081	cccgctcat	tggtcg	cgccattac	gcgcggg	aaagaagtcagctcggtatggcgcaaa	1140
Qy	1141	ggcgaagaa	atacaat	caaaagct	tgggcaac	gcgcgcgagcatataccggcgcatca	1200
Db	1141	ggcgaagaa	atacaat	caaaagct	tgggcaac	gcgcgcgagcatataccggcgcatca	1200
Qy	1201	tccaaaa	cagatga	agcccttc	acccgc	aacagcaatccggtacgttctcggtcgaatcc	1260
Db	1201	tccaaaa	cagatga	agcccttc	acccgc	aacagcaatccggtacgttctcggtcgaatcc	1260
Qy	1261	gaaggcata	tacccttc	agacata	taccgc	agcagcggaacacccctcgctcgctacgg	1320
Db	1261	gaaggcata	tacccttc	agacata	taccgc	agcagcggaacacccctcgctcgctacgg	1320
Qy	1321	gtttccg	cgcgcg	gagacgc	gcaggtt	aaagcagcgcgacgaataatcttagccgcagg	1380
Db	1321	gtttccg	c				

XX	AC	Z54511;
XX	DT	21-MAR-2000 (first entry)
XX	DE	Neisseria meningitidis ORF 986 partial DNA sequence SEQ ID NO:2969.
XX	DE	Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW	KW	antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW	KW	antibacterial; gene therapy; ds.
XX	OS	Neisseria meningitidis.
XX	OS	Neisseria meningitidis.
XX	PN	W09957280-A2.
XX	PN	11-NOV-1999.
XX	PD	
XX	PF	30-APR-1999; 99WO-US09346.
XX	PR	01-MAY-1998; 98US-0083758.
XX	PR	31-JUL-1998; 98US-0094869.
XX	PR	02-SEP-1998; 98US-0098994.
XX	PR	02-SEP-1998; 98US-0099062.
XX	PR	09-OCT-1998; 98US-0103749.
XX	PR	09-OCT-1998; 98US-0103794.
XX	PR	09-OCT-1998; 98US-0103796.
XX	PR	25-FEB-1999; 99US-0121528.
XX	PA	(CHIR) CHIRON CORP.
XX	PA	(GENO-) INST GENOMIC RES.
XX	PI	Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
XX	PI	Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
XX	PI	Tettelin H, Venter JC;
XX	PI	WPI; 2000-062150/05.
XX	DR	P-PSDB; Y75749.
XX	DR	Novel Neisserial polypeptides predicted to be useful antigens for
XX	PT	vaccines and diagnostics -
XX	PS	Claim 7; Page 1390; 1453pp; English.
XX	CC	Z53015 to Z54536, Z54577 to Z54615, and Y74253 to Y75941 represent
XX	CC	novel Neisseria meningitis and N. gonorrhoeae polynucleotides and
XX	CC	polypeptides. Z54537 to Z54576 and Z54616 to Z55473 represent PCR
XX	CC	primers used in the exemplification of the present invention. The
XX	CC	polypeptides, the polynucleotides, antibodies and compositions of the
XX	CC	invention can be used as vaccines, as diagnostic reagents, and as
XX	CC	immunogenic compositions. The polypeptides can be used in the
XX	CC	manufacture of medicaments for treating or preventing infection due to
XX	CC	Neisserial bacteria (e.g. meningitis and septicemia), to detect the
XX	CC	presence of Neisseria bacteria or to raise antibodies. They may also
XX	CC	be used to screen for agonists or antagonists, which may themselves
XX	CC	have use as antibacterial agents. The polynucleotides of the invention
XX	CC	may also be used in gene therapy protocols.
XX	SO	Sequence 1500 BP; 386 A; 469 C; 382 G; 263 T; 0 other;

```

Query Match      94.3%; Score 1315.8; DB 21; Length 1500;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 1352; Conservative 0; Mismatches 42; Indels 1; Gaps 1
y 1 ggttcaaaaaataccaataactcgtctttggcgggcaactgtgtccgccttctgtgcaggc 60
|||||
b 1 ggttcaaaaaataccaataacctcgtctttggcagcaactgtgtcagcctcgtgtgcaggc 60
|||||
y 61 tgcgaaaggcgagcgacttttctcgtgcgacaaaaaagaagcatcctctctagaacgc 120
|||||
b 61 tgcgaaaggcgagcgacttctctcgttgcgacaaaaaagaagcatcctctctagaacgc 120
|||||
y 121 atcgacacacacaaagacacagcagcatgtcagtatgctctgccgcactttgcccaactg 180

```

RESULT 6

REF: 254511

254511
ID 254511 standard; DNA; 1500 BP.

|||||
Db 121 atcgaaacacaaagacgagcgagcgtagtgctgctgcccgaactttgcccaactg 180
QY 181 gttcaaaagcaagcccgagtcgtcaatttcaggcagccccccgccccgcgaccccaa 240
Db 181 gttcaaaagtcagcgccgagtcgtagtcaatttcaggcagccccccgccccgcgaccccaa 240
QY 241 aacggcagcgcaatgcgcgaacacgattccgacccgcttgccgacgagacccgcttcac 300
Db 241 aacggcagcgcaatgcgcgaacacgattccgacccgcttgccgacgagacccgcttcac 300
QY 301 gaatttttcaaacgcctctgcccgaacatgcgccgaatccccccaaagaagacagatgac 360
Db 301 gaatttttcaaacgcctctgcccgaatgcgccgaatccccccaaagaagacagatgac 360
QY 361 ggcggattgaactcggctcggttcggttcacatcagcaaaaacggctacatcctgaccaat 420
Db 361 ggcggattgaactcggctcggttcggttcacatcagcaaaaacggctacatcctgaccaat 420
QY 421 acccacgctgttcggctgattgggcagtatcaaatccctgctcaacgacgaagcgcaatat 480
Db 421 acccacgctgttcggctgattgggcagtatcaaatccctgctcaacgacgaagcgcaatat 480
QY 481 accgccaactcatcgttctggatgtccaatccgatgtcgccctctgaaaaatcgacgca 540
Db 481 accgccaactcatcgttctggatgtccaatccgatgtcgccctctgaaaaatcgacgca 540
QY 541 acggaagatcaccgctcgtcataatcggcaatcccaatcccaaaattgaaacccggcgaatgg 600
Db 541 acggaagatcaccgctcgtcataatcggcaatcccaatcccaaaattgaaacccggcgaatgg 600
QY 601 gtcgctgccatcgccgccccttcggcttgacacagcgtgacccgcggaatcgctgccc 660
Db 601 gtcgcgccatcgccgccccttcggcttgacacagcgtgacccgcggaatcgctgccc 660
QY 661 gccaaagcgagaacccgtgccaaagaaagctacacacaccccttcataccaaacgcagcttgcc 720
Db 661 gccaaagcgagaacccgtgccaaagaaagctacacacaccccttcataccaaacgcagcttgcc 720
QY 721 atcaatccgggcaatttcggcgcccctgtgtcaacttaaaaggacaggtcgctcgccatc 780
Db 721 atcaatccgggcaactcgccgcccctgtgtcaacttaaaaggacaggtcgctcgccatc 780
QY 781 aatcgcgaatatatacagccgagcgggatttcattgggcattctcttggccatcccgatt 840
Db 781 aatcgcgaatatatacagccgagcgggatttcattgggcattctcttggccatcccgatt 840
QY 841 gacgttcccatgaatgtcgccgaacagctgaaacacacccgcgaagtcaccaacgcggacaa 900
Db 841 gacgttcccatgaatgtcgccgaacagctgaaacacacccgcgaagtcaccaacgcggacaa 900
QY 901 ctggcgctgattattcaggaagtatcctacggttttgccacagtcgtctggctggatataa 960
Db 901 ctggcgctgattattcaggaagtatcctacggttttgccacaaatcgttgggttggacaaa 960
QY 961 gccagcgcgcatgattgccaataatcttcccgagcccccgcagaaacgtgcggccctg 1020
Db 961 gccagcgcgcatgattgccaataatcttcccgagcccccgcagaaacgtgcggccctg 1020
QY 1021 caggcgcgacatcgtcctcagctcgacgagcgaggaataacgtttcttcggcgacatt 1080
Db 1021 caggcgcgacatcgtcctcagctcgacgagcgaggaataacgtttcttcggcgacatt 1080
QY 1081 ccggtcatggtcgcgccattacccggggaagaagtcagcctcggtcggtatggcgcaaa 1140
Db 1081 ccggttatggtcgcgccattacccggggaagaagtcagcctcggtcggtatggcgcaaa 1140
QY 1141 ggcgaagaataacaaatacaagccaagctgggcaacgcgcgcgagcatatccggcgcatca 1200
Db 1141 ggcgaagaataacaaatacaagccaagctgggcaacgcgcgcgagcatatccggcgcatca 1200
QY 1201 tccaaacagatgaagcccctcacgaaacagcaaatccgtagcttctcgtgaaatcc 1260
|||||

Db 1201 tccaaacagatgaagccccctacacagcaacacccgtagcttctcgtgtaatcc 1260
QY 1261 gcaggcatcacccttcagacacataccgacagcgcgcaaacaccctcgtcgtcqlacgg 1320
Db 1261 gtcaggcatcacccttcagacacataccgacagcgcgcaaacaccctcgtcgtcqlacgg 1320
QY 1321 gtttcgacgcgcaacgcgaggttgaagcgcgcgacgaaatccctagccgctcagg 1380
Db 1321 gtttcgacgcgcaacgcgaggttgaagcgcgcgacgaaatccctagccgctcagg 1380
QY 1381 gcaagtcccccgtaaa 1395
Db 1380 gcaagtcccccgtaaa 1394
RESULT 7
ID Z54512 standard; DNA; 1500 BP.
XX
AC Z54512;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 986 partial DNA sequence SEQ ID NO:2971.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy; ds.
XX
OS Neisseria meningitidis.
XX
PN WO9557280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
DR
WPI: 2000-062150/05.
DR P-PSDB; Y75750.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
XX vaccines and diagnostics .
PS Claim 7; Page 1392; 1453pp; English.
XX
CC Z53015 to Z54536, Z54577 to Z54615, and Y74253 to Y75941 represent
CC novel Neisseria meningitidis and N. gonorrhoeae polynucleotides and
CC polypeptides. Z54537 to Z54576 and Z54616 to Z5473 represent PCR
CC primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention

CC may also be used in gene therapy protocols.

XX
SO
Sequence 1500 BP; 382 A; 469 C; 382 G; 265 T; 2 other;

Query Match 94.2%; Score 1313.8; DB 21; Length 1500;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 1350: Conservative 0; Mismatches 44; Indels 1;

1:

[illegible]

Qy	961	gcacgaggcgcatgtgatggccaaaatccttcccggcagcccgacgaacgtgccgacctg	1020
Db	961		
Qy	1021	caggcgggcgacatcgtcttcagctcgacgcggcgagaaatacgtttcttcggcgacctt	1080
Db	1021		
Qy	1081	ccccgtcatggtcggcgccatttacgcggggaaaaagaagtcaaccttcggcggtatggcgcaaa	1140
Db	1081	ccccgttatggtcggcgccatttacgcggggaaaaagaagtcaaccttcggcggtatggcgcaaa	1140
Qy	1141	ggcgagaagaatacacatacaaacccaagctggggcaaccgcgcgagcatacccgcgcatca	1200
Db	1141		
Qy	1201	tccaaacagatgaagcccccttacaccgaacagcaatccctgttcgttcggtcgnaatcc	1260
Db	1201	tccaaacagatgaagcccccttacaccgaacagcaatccctgttcgttcggtcgnaatcc	1260
Qy	1261	gcaggcatcacccttcagacacataccgacagcgcaaacactcgttcgttcgtaacgg	1320
Db	1261	gcaggcatcacccttcagacacataccgacagcgcaaacactcgttcgttcgtaacgg	1320
Qy	1321	gtttccgacgvcgcagaaacgcgaggctttaaggcagcgcgcagaaatcctcagccgtcagg	1380
Db	1321	gtttccgacgvcgcagaaacgcgaggcttgaggcgcgcgcagaaatccttcgcgc-tc-gg	1379
Qy	1381	gcaagtccccgtcaa	1395
Db	1380	qcaagtccccgtcaa	1394

RESULT 8

7.333061

ID Z33306 standard: DNA; 1500 BP.

XXXXXX

AC 233306:

22 XX

DT 21-FEB-2000 (first entry)

XX *Neisseria meningitidis* strain ATCC 13090 BASB013 nucleotide sequence.

XX *Neisseria meningitidis*: diagnosis; infection; vaccine;
KW *Neisseria meningitidis*: BASB013; diagnosis; infection; vaccine;

KW antibiotic; upper respiratory tract infection; bacteraemia; meningitis

KW invasive bacterial disease; antibacterial; ss.

Neisseria meningitidis.

2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526

PN W09955872-A1.

XXXXXX

04-NOV-1999.

2 X
2 X

PF 20-APR-1999; 99WO-EP02765.

XX

PR 23-APR-1998; 98GB-0008734.

XX PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX

PI Ruelle J;

XX
XX

DR WPI; 2000-052809/04.

DR P-PsDB; Y52994.

xx Novel polynucleotides and polypeptides from *Neisseria meningitidis* used
pt to prepare vaccines against bacterial infections -

XX
XX

PS Claim 12:

XX
 22

CC The present sequence encodes a BASB013 polypeptide isolated from
CC *Neisseria meningitidis*. BASB013 polynucleotides and polypeptides may be
CC employed as research reagents and material for the discovery of
CC treatments and diagnostics for diseases, particularly human diseases.

CC They can be used for diagnosis of disease, staging of disease, or
CC determining response of an infectious organism to drugs. The
CC polynucleotides may be used as a source for hybridisation probes, and
CC for screening of genetic mutations, serotype, organism or strain
CC identification, identification of mutation in BASB013 sequences, and as
CC components of arrays which are useful for diagnostic and prognostic
CC purposes. The polypeptides can be used to produce antibodies. The
CC polypeptides can also be used in vaccine formulations, and to identify
CC agonists and antagonists. The polypeptides, antibodies, agonists and
CC antagonists (which are bacteriostatic) are used for the treatment and
CC prevention of diseases such as upper respiratory tract infection,
CC invasive bacterial diseases such as bacteraemia and meningitis, and for
CC the development and screening of antibacterial drugs. They are also used
CC in the prevention of adhesion of bacteria to eukaryotic matrix proteins
CC on in-dwelling devices, or to extracellular proteins on wounds, and to
CC thus prevent tissue damage and/or block the normal progression of
CC pathogenesis in infections initiated other than by the implantation of
CC in-dwelling devices or by other surgical techniques.
XX
SQ Sequence 1500 BP; 370 A; 469 C; 392 G; 269 T; 0 other;

Query Match 90.5%; Score 1263; DB 21; Length 1500;
Best Local Similarity 94.6%; Pred. No. 0;
Matches 1319; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

Qy 1 gtgttcaaaaaataccataactcttctgttggcgcactgtgtgcgccttgcgtggcaggc 60
Db 1 gtgttcaaaaaataccataactcttctgttggcgcactgtgtgcgccttgcgtggcaggc 60
Qy 61 tgcgaaaggcaggcagcttttctgttgcgcgcaaaaaaaggagcatccttctgtagaacgc 120
Db 61 tgcgaaaggcaggcagcttttctgttgcgcgcaaaaaaaggagcatccttctgtagaacgc 120
Qy 121 atcgaacacacaaagacgacgagcttcaatgtcagtctgctccgacactttgcccaactg 180
Db 121 atcgaacacacaaagacgacgagcttcaatgtcagtctgctccgacactttgtccaaactg 180
Qy 181 gtccaaagcgaagcccgccagctgcataattcaggcagcccgcccgcgaccccaa 240
Db 181 gtccaaagcgaagcccgccagctgcataattcaggcagcccgcccgcgaccccaa 240
Qy 241 aacggcagcgcaatgcgcgaaacccgattccgaccgcttgcgcgacagcaccgcttcac 300
Db 241 aacggcagcgcaatgcgcgaaacccgattccgaccgcttgcgcgacagcaccgcttcac 300
Qy 301 gaattttcaaacgctcgtcccgacatgccggaatcccgaaatcccccagaagaagcagatgac 360
Db 301 gaattttcaaacgctcgtcccgacatgccggaatcccgaaatcccccagaagaagcagatgac 360
Qy 361 ggcggattgaacttcggttcggttcacatcagcaaaacggctacatcctcgaccaat 420
Db 361 ggcggattgaacttcggttcggttcacatcagcaaaacggctacatccttcgaccaat 420
Qy 421 accacgctgttgcggtatggcgagatcaaaagtctcgtcacaagcagaagcggaatat 480
Db 421 acgcaactcgttacggcagtgagcagatcaaaagtctcgtcacaagcagaagcggaatat 480
Qy 481 accgcaaaactcactcgttcgagatgccaatcccgatgtcgcccttctgaaatcgcgcga 540
Db 481 accgcaaaactcactcgttcgagatgccaatcccgatgtcgcccttctgaaatcgcgcga 540
Qy 541 acggaagagctaccgctgcataaatcggcaatcccaaaaatttgaaacccggcggaatgg 600
Db 541 acggaagagctgcccgctgcataaatcggcaatcccaaaaatttgaaacccggcggaatgg 600
Qy 601 gtcgctgcacatggcgcccttgcgttttgacaacagcgtgacgcgcggtatcgtgtcc 660
Db 601 gtcgccgcacatggcgcccttgcgttttgacaacagcgtgacgcgcggtatcgtgtcc 660
Qy 661 gccaaaggcagaagcctgcccaacagaagctacacaccccttcacaaacccagcttgc 720
Db 661 gccaaaggcagaagcctgcccaacagaagctacacaccccttcacaaacccagcttgc 720

Qy 721 atcaatccgggcaattccggcgccgctgttcaactttaaaaggacaggtctcggcattc 780
Db 721 atcaatccgggcaattccggcgccgctgttcaactttaaaaggacaggtctcggcattc 780
Qy 781 aattcgcaaatatcacgccgcagcgcgaggtatcattgggcatctcttgcacatcccgatt 840
Db 781 aactcgcaaatatcacgccgcagcgcgaggtatcattgggcatctcttgcacatcccgatt 840
Qy 841 gacgttccactgaatgcgcgcaaacactgaaacacacgcgcaaaatccaacgcgacaa 900
Db 841 gacgttccactgaatgcgcgcaaacactgaaacacacgcgcaaaatccaacgcgacaa 900
Qy 901 ctggcgctgattatcagggaagtatcctcactggtttggcacagctctcgtgtcgatataa 960
Db 901 ctggcgctgattatcagggaagtatcctcactggtttggcacagctctcgtgttgacaaa 960
Qy 961 gccagcgcgcatgtattgcacaaatccttcccgcgacgccccgcgacgagcgtgcgcctg 1020
Db 961 gccgcgcgcatgtattgcacaaatccttcccgcgacgccccgcgacgagcgtgcgcctg 1020
Qy 1021 caggcgcgcaacatcctcactgacgctgcagcgcgaggaatacgttcttcggcgcaactt 1080
Db 1021 caggcgcgcaacatcctcactgacgctgcagcgcgaggaatacgttcttcggcgcaactt 1080
Qy 1081 ccgctcatgttcggcgccattacgcgcgggaaaaaagtcagcctcggcggtatggcgcaaa 1140
Db 1081 ccgctcatgttcggcgccattacgcgcgggaaaaaagtcagcctcggcggtatggcgcaaa 1140
Qy 1141 ggcgaaagaatacaaaataaaagccaaagctgggcaacgcgcgcgagcatcaccggcgcatca 1200
Db 1141 ggtaaaggaatacaacgttgcgctcaaaactggcgaatgcttccgaaacaaacccggttccctg 1200
Qy 1201 tccaaacacagatgaagccctctacacgcaacagacaatccgtacgcttctcgttcaaatcc 1260
Db 1201 tccgagcggaacaaagccctctatgcggaacacaaatccggaacttctcgttcaaatcc 1260
Qy 1261 gcaggcattacccttcagacacataccgacagcgcggaacacccctcgtctcgtacgg 1320
Db 1261 gcaggcattacccttcagacacataccgacagcgcggaacacccctcgtctcgtacgg 1320
Qy 1321 gtttcgacgcgcgacgaacgcgaggtttaaaggcacggcgacgcaaaatcctcagcgtcagg 1380
Db 1321 gtttcgggggcgagacgcgaggttggcgcgaggttggcgcgagcgaatccttgcgctc-gg 1379
Qy 1381 gcaagtcgcccgtaaa 1395
Db 1380 gcaagtcgcccgtaaa 1394
RESULT 9
233305
ID 233305 standard; DNA; 1500 BP.
XX
AC 233305;
XX
DT 21-FEB-2000 (first entry)
XX
DE Neisseria meningitidis strain ATCC 13090 BASB013 nucleotide sequence.
XX
KW Neisseria meningitidis; BASB013; diagnosis; infection; vaccine;
KW antibiotic; upper respiratory tract infection; bacteraemia; meningitis;
KW invasive bacterial disease; antibacterial; ss.
XX
OS Neisseria meningitidis.
XX
PN W0955872-A1.
XX
PD 04-NOV-1999.
XX
PF 20-APR-1999; 99WO-EP02765.
XX
PR 23-APR-1998; 98GB-0008734.

[illegible]

DT	24-JUL-2000	(first entry)	
XX	NGSP polypeptide coding sequence.		
XX			
XX	NGSP: polypeptide; peptide; vaccine; immune response; antibody;		
KW	cellular matrix; adherence domain; ligand; detection; diagnosis;		
KW	screening; probe; primer; prophylaxis; therapy: ss.		
XX			
OS	Neisseria gonorrhoeae.		
XX			
XX	Key	Location/Qualifiers	
FT	CDS	1..1242	
FT	/*tag= a		
FT	/product= NGSP polypeptide		
XX	WO200012133-A1.		
XX			
PD	09-MAR-2000.		
XX			
XX	01-SEP-1999;	99WO-US20070.	
XX			
PR	01-SEP-1998;	98US-0098685.	
XX			
PA	(ANTE-) ANTEX BIOLOGICS INC.		
XX			
PI	Jackson WJ,	Harris AM;	
XX			
DR	WPI: 2000-237782/20.		
DR	P-PSDB; Y83151.		
XX			
PT	Non-cytosolic NGSP polypeptide and polynucleotide sequence from		
PT	Neisseria useful for diagnosis, prevention or treatment of Neisseria		
PT	infections		
XX			
PS	Claim 38; Page 62-63; 68pp; English.		
XX			
CC	The NGSP polypeptide of N. gonorrhoeae has conserved Arg-Gly-Asp and		
CC	Arg-Gly-Asn motifs near the C-terminus which function as adherence		
CC	domains for extracellular matrix proteins. Using the NGSP polypeptide		
CC	as a vaccine produces antibodies which inhibit binding of N.		
CC	gonorrhoeae to the host's cellular matrix reducing attachment and/or		
CC	subsequent invasion. The NGSP polypeptide and its peptide fragments		
CC	can be used to immunise an animal and produce an immune response.		
CC	They can also be used as ligands to detect antibodies elicited in		
CC	response to Neisseria infections and also as antigens or immunogens		
CC	for inducing Neisseria-specific antibodies which are useful in		
CC	immunoassays to detect Neisseria in biological specimens. Nucleotides		
CC	encoding NGSP or its fragments can be used as probes to identify		
CC	Neisseria in biological specimens by hybridization or polymerase		
CC	chain reaction amplification. The NGSP polypeptide can also be used		
CC	in screening assays to identify agents and compounds which useful as		
CC	diagnostic, prophylactic or therapeutic agents against Neisseria		
CC	infection.		
XX			
SQ	Sequence 1242 BP: 316 A; 400 C; 312 G; 214 T; 0 other;		
Query Match			
Best Local Similarity 89.0%; Score 1242; DB 21; Length 1242;			
Matches 1242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	154	atgctgtgccgactttgcccaactgtttcaaacggaagcccgagctgtcaattt	213
DB	1	atgtgtgtcccgactttgcccaactgtttcaaacggaagcccgagctgtcaattt	60
QY	214	caggcagcccccgcgcgcacccaaacgagcggcggaatgccgaacccgattccgac	273
DB	61	caggcagcccccgcgcgcacccaaacgagcggcggaatgccgaacccgattccgac	120
QY	274	ccgcttgcgcagcagccctttctacgaatttttcaaacgctgtcccgcacatgcc	333
DB	121	ccgcttgcgcagcagccctttctacgaatttttcaaacgctgtcccgcacatgcc	180

```

RESULT 11
251533
ID 251533 standard; DNA; 1347 BP.
XX
AC 251533;
XX
DT 03-JUL-2000 (first entry)
XX
DE Neisseria meningitidis NMASP protein-1 encoding DNA.
XX
KW NMASP; non-cytosolic; antibacterial; antiinflammatory; cytotoxic;
KW anti-NMASP antibody; vaccine; diagnosis; therapy; prophylaxis;
KW Neisserial infection; meningitis; septicaemia; ds.
XX
OS Neisseria meningitidis.
XX
FH Key Location/Qualifiers
FT 1..1347
CDS
FT /*tag= a
FT /product= "Neisseria meningitidis protein"
FT /transl_except= (pos:499..501, aa: Xaa)
FT /note= "Xaa is unknown"
XX
PN W0200012535-A2.
XX
XX 09-MAR-2000.
XX
PF 01-SEP-1999; 99WO-US19663.
XX
PR 01-SEP-1998; 98US-0098685.
XX
PA (ANTE-) ANTEX BIOLOGICS INC.
XX
PI Jackson WJ, Harris AM;
XX
XX WPI; 2000-256581/22.
DR P-PSDB; Y70409.
XX
XX Neisseria meningitidis NMASP polypeptide, nucleotide sequences and
PT antibodies, useful in vaccines against infection
XX
PS Claim 41; Page 63; 75pp; English.
XX
XX The present sequence is a DNA encoding Neisseria meningitidis NMASP
XX protein. NMASP is a non-cytosolic protein, with antibacterial and
XX antiinflammatory activity. It shows sequence similarity to E. coli
XX Degrp (HtrA) protein. NMASP proteins can be used as ligands to detect
XX antibodies elicited in response to N. meningitidis infections. Cytotoxic
XX anti-NMASP antibodies can be used as vaccines. NMASP proteins and DNA
XX may be used for diagnosis, therapy or prophylaxis of Neisserial
XX infections such as, bacterial meningitis and septicaemia.
XX Note: The protein sequence represented in SEQ ID NO:2 of the
XX specification is erroneous"
XX
SQ Sequence 1347 BP; 338 A; 428 C; 344 G; 236 T; 1 other;

Query Match 84.3%; Score 1176.2; DB 21; Length 1347;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1207; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

QY 154 atgtgtgtccgactttgcccaactgtttcaaaagcgaagcccgccgagtcgcaatt 213
DB 1 atgtgtgtccgactttgtccaaactgtttcaaaagcgaagcccgccgagtcgcaatt 60
QY 214 caggcagcccccccgccgaccccaaaacgagcggcgaatgcgaaacccgattccgac 273
DB 61 caggcagcccccccgccgaccccaaaacgagcgaatgcgaaacccgattccgac 120
QY 274 ccgcttgcgacgacccgtttcacgaattttcaaacgctctgctccgaaatgcc 333
DB 121 ccgcttgcgacgacccgtttcacgaattttcaaacgctctgctccgaaatgcc 180

```

```

QY 334 gaaatcccccagaagaacagatgacggcggaattgaacttcggttcgggttcacatc 393
DB 181 gaaatcccccagaagaacagatgacggcggaattgaacttcggttcgggttcacatc 240
QY 394 agcaaaacggtacatctctgaccataccacgctogtttcoggtatggcgagtatacaa 453
DB 241 agcaaaacggtatattctgaccaaatacgcacgctcttaccgcatggcgagtatacaa 300
QY 454 gtctgtctcaacgacgacgcaataataccgcaaacctatcggttcggtatgctcaatcc 513
DB 301 gtctgtctcaacgacgacgcaataataccgcaaacctatcggttcggtatgctcaatcc 360
QY 514 gatgtcgcccttctgaaatcgacgaacggaagagctacccgtcgtcaaaatcggaat 573
DB 361 gatgtcgcccttctgaaatcgacgaacggaagagctcgcgtcgtcaaaatcggaat 420
QY 574 ccaaaaaattgaacccggcggaatgggtcgctgcctatcgctgcgccttcgggtttgac 633
DB 421 ccaaaagatttgaacccggcggaatgggtcgctgcgccttcgggttcgac 480
QY 634 aacagcgtacccgacgcatcgtgtcgccaaagcagaagcctgcccacgaagaagctac 693
DB 481 aacagcgtacccgacgcatcgtgtcgccaaagcagaagcctgcccacgaagaagctac 540
QY 694 acaccttcatccaaacccgacgttgcctcaatccgggcaattccggcgcccgctgttc 753
DB 541 acaccttcatccaaacccgacgttgcctcaatccgggcaattccggcgcccgctgttc 600
QY 754 aacttaaaaggacagtgctgcgcataatctgcgaataatacagccgagcgggattc 813
DB 601 aacttaaaaggacagtgctgcgcataatctgcgaataatacagccgagcgggattc 660
QY 814 atggcactctcttgcctcccgatgacgttgcctgaatgcgcgaacagctgaaa 873
DB 661 atggcactctcttgcctcccgatgacgttgcctgaatgcgcgaacagctgaaa 720
QY 874 aacacggcaaatcccaacggaacactggcgctgattattcagggaagtacctacggt 933
DB 721 aacacggcaaatcccaacggaacactggcgctgattattcagggaagtacctacggt 780
QY 934 ttggcagctcgtgtgtgtgataaaagccagcggcgctgattggtccaaatccttccc 993
DB 781 ttggcacaatctcgtgtgtgtgataaaagccgagcggcgctgattggtccaaatccttccc 840
QY 994 ggacgccccgacgaacgctgcgcctgcagcggcggaacatcgctcctcagctcagcgc 1053
DB 841 ggacgccccgacgaacgctgcgcctgcagcggcggaacatcgctcctcagctcagcgc 900
QY 1054 ggagaatacgtttcttcgcggaaccttcctgcgtatggtcgcgcgcattacgcgcggaaa 1113
DB 901 ggagaatacgtttcttcgcggaaccttcctgcgtatggtcgcgcgcattacgcgcggaaa 960
QY 1114 gaagtacgctcgtgtatggcgcaaaagggaagaatacaaatcaaaagcgaagctgggc 1173
DB 961 gaagtacgctcgtgtatggcgcaaaagggaagaatacaaatcaaaagcgaagctgggc 1020
QY 1174 aacgcccgcgagcagatccgcgcatactccaaacagatgaagccctctacacgaacag 1233
DB 1021 aacgcccgcgagcagatccgcgcatactccaaacagatgaagccctctacacgaacag 1080
QY 1234 caatcgggtacgttctcgttcgaatccgcgaggtattacccttcagacacataccacagc 1293
DB 1081 caatcgggtacgttctcgttcgaatccgcgaggtattacccttcagacacataccacagc 1140
QY 1294 agcggaaaacacactcgtcgtacggtttccacgacgagcagacgcagcagcttaagg 1353
DB 1141 agcggggagacacactcgtcgtacggtttccacgacgagcagcagcagcagcttaagg 1200
QY 1354 cagcgcgacgaatactcagcgtcagggcgaagtcccgctcaa 1395
DB 1201 cgcggcgacgaatacttccgctc-gggcaagtcctccgcaa 1241

```

QY	122	tcgaacacacacaaagaagacgagcgagtgatgtagtattgctgtgtcccgacttttgcccact	
Db	122	tcgaacacacacaaagaagacgagcgagtcagtagtctgtgccgactttgcccact	
QY	182	tcaaaagcgaaagcccgcgagtcgtcgaattatcagcgagccccgcgcgcgcaaccacaa	241
Db	182	tcaaaagtgaaggtcgcgcgagtcgtcgaattatcagcgagccccgcgcgcgcaaccacaa	241
QY	242	acggcagcggaattccgaaaccgattccgaaccgccttgccgaacgacccctgtctacg	301
Db	242	acggcagcggaattccgaaaccgattccgaaccgccttgccgaacgacccctgtctacg	301
QY	302	aattttcaaacgctctcccgaaacatcccgaaatccccaaagaagaagcagatgacg	361
Db	302	aattttcaaacgctctcccgaaatctccgaaatccccaaagaagaagcagatgacg	361
QY	362	gcggaattgaacttcggttcgggtctcatcatcagcaaaaaacggctacatctctgaccaata	421
Db	362	gcggaattgaacttcggttcgggtctcatcatcagcaaaagcggctacatctctgaccaata	421
QY	422	ccacgttcgttgcggtatggcgagtatcaaaagtcctgctcaacgacaaacgcgaataata	481
Db	422	ccacgttcgttaccggcattggcgagtatcaaaagtcctgctcaacgacaaacgcgaataata	481
QY	482	ccgcacaaactcatcggttcgggttcggaatccgaatccgattcgcccttctgaaatactgacgcaa	541
Db	482	ccgcacaaactcatcggttcgggttcggaatccgaatccgattcgcccttctgaaatactgacgcaa	541
QY	542	cggaaagactaccgttcgttcaaaaaatgggcaatccccaaaaatttgaaacccgggcgaatcggg	601
Db	542	cggaaagactgccttcgttcaaaaaatgggcaatccccaaagatttgaaacccgggcgaatcggg	601
QY	602	tgcgtgcacatcggcgcgcccttcgctcttgaaacagcgtgacccgcgcgcgcatcgctgccg	661
Db	602	tgcgcgcctatcggcgcgcccttcgctctcgaaacacgctgacccgcgcgcgcatcgctgccg	661
QY	662	ccaaaggcagaagcctgcccacgaaagctacacacccctctatccaaacgcgacgcttgcca	721
Db	662	ccaaaggcagaagcctgcccacgaaagctacacacccctctatccaaacgcgacgcttgcca	721
QY	722	tcaatccgggcaattccggcgcccgctgttcaactttaaaggacaggttcgttcgggcatca	781
Db	722	tcaatccgggcaattccggcgcccgctgttcaactttaaaggacaggttcgttcgggcatca	781
QY	782	attcgcaaatatacagcgcgacgcgcgagattcatgggcattcctcttgccatccccgattg	841
Db	782	actcgcaaatatacagcgcgacgcgcgagattcatgggcatttctcttcgcacatccccgattg	841
QY	842	acggtgcataaattgcgggaacagctgaataaacacccggcaaaagtcacacgcggacacac	901
Db	842	acggtgcataaattgcgggaacagctgaataaacacccggcaaaagtcacacgcggacacac	901
QY	902	tggcgctgattattcaggaagtatcctaaggcttggcacagtcggttcgggtctggataaag	961
Db	902	tggcgctgattattcaggaagtatcctaaggcttggcacacatcggttcgggtctggataaag	961
QY	962	ccagcgcgcgcattgattgccaaaaatccttcccgggaagccccgcgacagctgcgcgcctgc	1021
Db	962	ccgcgcgcgcgcatgattgccaaaaatccttcccgggaagccccgcgacagctgcgcgcctgc	1021
QY	1022	aggcgcgcgcacatgctctcagcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgccttc	1081
Db	1022	aggcgcgcgcacatgctctcagcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgccttc	1081
QY	1082	ccgctatggtcggcgccattacgcgcggga	1110
Db	1082	ccgctatggtcggcgccattacgcgcggga	1110

ID	X98241 standard	DNA	1436 BP
X98241			

Fri Mar 23 07:26:46 2001

Db 1147 agccaggttctgcctaattctctccgtgcaaaagcgggcattaaagcgggtgatgtgatc 1206
Qy 1039 ctacgctcgacggcgagaaatacgttcttccggcgaccttccgtcatgtcgggcgcc 1098
Db 1207 acctcaactgaacgggtaagccgatacagcagctttgccgactgcgtcaggtgggtact 1266
Qy 1099 attacgcccgggaaagaaagtcagcctcggcgtatggcgcnaaaggcgaagaaatcacaaatc 1158
Db 1267 atgccggtaggcagcaaaactgacctgggttactgcgacggttaagcaggttaacgtg 1326
Qy 1159 aaagccaaagctgggcaacgcgcgcagcat 1188
Db 1327 aacctggaactgccgcagagcagccagaat 1356

Search completed: March 22, 2001, 10:17:11
Job time: 25016 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2001, 02:51:15 ; Search time 58.35 Seconds
(without alignments)
3852.933 Million cell updates/sec

Title: US-09-388-090-3
Perfect score: 1395
Sequence: 1 gtttcaaaaataccaata.....tcaggggcaagtcctccgcaaa 1395

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 280836 seqs, 80580151 residues
Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, NA: *
1: /cgn2.6/prodata/2/ina/5a_COMB.seq: *
2: /cgn2.6/prodata/2/ina/5a_COMB.seq: *
3: /cgn2.6/prodata/2/ina/5_COMB.seq: *
4: /cgn2.6/prodata/2/ina/PCTUS_COMB.seq: *
5: /cgn2.6/prodata/2/ina/backfiles1.seq: *

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	176.4	12.6	1854	3	US-08-923-454A-13
2	175.6	12.6	1980	1	US-08-350-741-1
3	175.6	12.6	1980	2	US-08-463-875A-1
4	154.2	11.1	1791	1	US-08-245-294-7
5	154.2	11.1	1791	1	US-08-474-499-7
6	154.2	11.1	1791	1	US-08-307-279A-7
7	154.2	11.1	1791	4	PCT-US95-06211-7
8	139.4	10.0	2894	1	US-08-278-091-1
9	139.4	10.0	2894	1	US-08-483-859-1
10	139.4	10.0	2894	1	US-08-472-173-1
11	139.4	10.0	2894	2	US-08-487-167-1
12	139.4	10.0	2894	2	US-08-482-816-1
13	139.4	10.0	2894	2	US-08-296-149-1
14	139.4	10.0	2894	2	US-08-801-499-1
15	139.4	10.0	2894	2	US-08-615-271-1
16	139.4	10.0	2894	3	US-09-074-660-1
17	139.4	10.0	2894	3	US-09-074-659-1
18	139.4	10.0	2894	3	US-09-106-468-1
19	139.4	10.0	2894	3	US-09-106-466A-1
20	139.4	10.0	2894	3	US-09-106-467-1
21	136.2	9.8	1611	1	US-08-485-569-1
22	136.2	9.8	1611	1	US-08-480-293-1
23	136.2	9.8	1611	2	US-07-903-079B-1
24	105	7.5	2036	3	US-08-923-454A-17
25	105	7.5	2205	3	US-08-888-077A-41
26	85.8	6.2	539	1	US-08-322-742-18
27	57.6	4.1	1835	3	US-08-923-454A-7
28	57.6	4.1	2187	3	US-08-923-454A-23

29	57.6	4.1	2187	3	US-08-923-454A-24
30	57.6	4.1	2187	3	US-08-923-454A-30
31	51.6	3.7	1503	3	US-08-923-454A-5
32	51.6	3.7	1787	3	US-08-923-454A-3
33	51.6	3.7	2144	3	US-08-923-454A-28
34	51.6	3.7	2951	3	US-08-923-454A-26
35	41	2.9	1187	1	US-08-440-856A-2
36	39.2	2.8	2414	5	5248599-1
37	37.2	2.7	2793	1	US-08-209-747-1
38	37.2	2.7	2793	1	US-08-458-298-1
39	36.8	2.6	716	2	US-08-211-718-1
40	36.8	2.6	1608	2	US-08-211-718-8
41	36.2	2.6	411	3	US-08-963-168C-1
42	36.2	2.6	411	3	US-08-963-168C-2
43	35.8	2.6	399	3	US-08-963-168C-5
44	35.8	2.6	2167	2	US-08-461-775-9
45	35.8	2.6	2167	3	US-09-031-606-9

ALIGNMENTS

RESULT 1
US-08-923-454A-13
; Sequence 13, Application US/08923454A
; Patent No. 6004794
; GENERAL INFORMATION:
; APPLICANT: Creasy, Caretha
; APPLICANT: Livi, George
; APPLICANT: Karan, Eric
; APPLICANT: Clinkenbeard, Helen
; APPLICANT: Browne, Michael
; APPLICANT: Southan, Christopher
; TITLE OF INVENTION: HUMAN SERINE PROTEASE
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,454A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/025436
; FILING DATE: 06-SEPT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: F50547
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1354 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:


```
QY 697 cccctcatcaaacacagctgtgcatcaatcccgggcaattcccgccgcccgtgttcaac 756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1064 AACTTTATTACAGACCGACCGCGGATTAACCGTGGTAACCTCCGGCGCGCTGGTGAAC 1123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 757 ttaaaagacagctgtctggcatcaattcgcaataataacagccgagcgcggtattcatg 816
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1124 CTGACGGGTGAGTGATCGGTATTAAACACCGCATTTCTGCGCGGACGGCGCAACATC 1183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 817 gcatctctttgcatcagctgagcttgcctgaatgtcgcgaacagctgaaaaac 876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1184 GGTATCGCTTCCTATPCCCACTAACAYGTCAAAACCTGACGTCCGACATGGTGGAA 1243
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 877 accggcaaaagtcacacgagcaactggcggtgattattcaggaaagtatcctacggttg 936
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1244 TACGCGCAGTGAACCGCGCGAATCGGCGATCATGGGACTGAGCTGAATTCGGAATTG 1303
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 937 gscagtcgttcgtctgtgataaagccagcgcgccattgattgccaaaaacccctccggc 996
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1304 GCGAAAGCGATGAAGTCGACGCCCGACGAGCGCGGTCTGTCAAGCCAGGTGATGCCGAAT 1363
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 997 agcccgacagacgctgcggcctgcagcgcgggcgagacatcgtcctcagcctcgacgcyga 1056
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1364 TCGTCCGCGGAAAGCGGGTATCAAGCGCGGGGATGTCAATTACCTCGCTGAACGGTAAA 1423
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1057 gaaatcgtttctccggcgaccttcccgctcatgtgctggcgccattacgcccgggaaagaa 1116
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1424 CCGATACGAGCTTTGCGGCGCTGCGCGCTCAGGTGCGCACATATGCCGTGCGGACGAAA 1483
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1117 gtacgctcgtgtatggcgaagcggaagaaatacaataaagccaagctg 1170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1484 ATCAGCCTCGGTCTGCTGTGAAGGTAAAGCGATTACGGTGAATCTGGAACCTG 1537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 3

US-08-463-875A-1

Sequence 1, Application US/08463875A

Patent No. 5980907

GENERAL INFORMATION:

APPLICANT: DOUGAN, Gordon

APPLICANT: CHARLES, Ian G.

APPLICANT: HORMACHE, Carlos E.

APPLICANT: JOHNSON, Kevin S.

APPLICANT: CHATFIELD, Steven N.

TITLE OF INVENTION: LIVE VACCINES

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESSES:

ADDRESSEE: NIXON and VANDERHYE PC

STREET: 8th FLOOR, 1100 NORTH GLEBE ROAD

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: USA

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/463,875A

FILING DATE: 05-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/340,741

FILING DATE: 07-DEC-1994

APPLICATION NUMBER: US 07/952,737

FILING DATE: 30-NOV-1992

APPLICATION NUMBER: GB 9007194.5

FILING DATE: 30-MAR-1990

APPLICATION NUMBER: PCI/GB91/00484

FILING DATE: 28-MAR-1991

ATTORNEY/AGENT INFORMATION:

NAME: WILSON, MARY J.

REGISTRATION NUMBER: 32,955

REFERENCE/DOCKET NUMBER: 117-158

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4000

TELEFAX: (703) 816-4100

TELEX: 200797 NIXN UR

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1980 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 395..1822

US-08-463-875A-1

Query Match 12.6%; Score 175.6; DB 2; Length 1980;

Best Local Similarity 52.6%; Pred. No. 2.6e-39;

Matches 407; Conservative 0; Mismatches 364; Indels 3; Gaps 1;

```
QY 400 aacggtcacatcctgaccaaataccacgctggttccggtatgggcagtatcaaatcctg 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 764 AAGGCTACGTCTGCACCAACACACCGTGGTTGATAACGCCAGCTGATTAAAGTACAG 823
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 460 ctcaacgacaagcggaatatataccgccaactcatcgttccggtgtccaatccgatgctc 519
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 824 CTTAGCGATGGCGCTAAATTCGATGCTAAAGTGTGGGCAAGATCCGCGTCTCTGATATC 883
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 520 gcccttctgaaatcgacgaacgaagagctaccgctcgtaaaatcggcaatcccaaaa 579
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 884 GCGTGATTCAAAATTCAGATCCGGAAGAACTGACGGCGATTAAAGTGGCGGACTCCGAC 943
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 580 aattgaaacccggcggaatgggtcgtgcccacgcgcgccccttcggtttgacaacagc 639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 944 GCGTGCCTGGTGGGGATTATACCGTCGCTATTGTGTAAACCCGTTTGGTCTGGGCGAAACG 1003
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 640 gtaccgccggcatcgtgtccgccaag ---gcagaagcctgcccacgaagactacaca 696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1004 GTGACGTGAGGTATCGTTTCGGCGCTTGGGGCGTACGGCGCTGAACGTAGAAAAATTCACGAG 1063
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 697 ccttcatcaaacacgcttggccatcaatccggaattccgcgcgcgcgctgttcaac 756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1064 AACTTTATTACACGCGCGCGGATTAACCGTGGTAACTCCGGGGCGCGCTGTGTGAAC 1123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 757 ttaaaagacagctgtcgtcgccatcaattcgcaaatatacacgccgagcgcggtattcatg 816
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1124 CTGAACGGTGAGCTGATCGGTATTAAACACCGCGATTCTTGGCGCGGACGGCGCAACATC 1183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 817 ggcattcctttgccatcccgattgacgttgcctgaatgtcgcgaacagctgaaaaac 876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1184 GGTATCGCTTCCTATPCCCACTAACATGGTGAACCACTGACCTCCGACATGGTGGAA 1243
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 877 accggcaaaagtcacacgagcggaactggcggtgattattcaggaaagtatcctacggttg 936
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1244 TACGCCACAGGTGAACCGCGCGAATCGGGATCATGGGACTGAGCTGAATTCGCAATTG 1303
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 937 gcacagtcgttcgtctgtgataaagccagcgcgccattgattgccaaaaacccctccggc 996
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1304 GCGAAAGCGATGAAGTCGACGCCCGACGAGCGCGGTTCGTCAAGCCAGGTGATGCCGAAT 1363
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 957 agccccgcaaacgtgcccgcctgcagcgcgagacatcgtcctcagctcagcgcgga 1056
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1364 TCGTCCGCGGCGAAAGCGGGTATCAAGCCGGGGATGTCAATTACCTGCTGAACGGTAAA 1423
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1057 gaaatacgtttctccggcgaccttcccgctcatgttccgccccattacgcccgggaaaaagaa 1116
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1424 CCGATACGAGCTTTGCGGCGCTGCGCGCTCAGGTGCGCACATATGCCGTGCGGACGAAA 1483
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1117 gtacgctcgtgtatggcgaagcggaagaaatacaataaagccaagctg 1170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1484 ATCAGCCTCGGTCTGCTGTGAAGGTAAAGCGATTACGGTGAATCTGGAACCTG 1537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 4
US-08-245-294-7
: Sequence 7, Application US/08245294
: Patent No. 5644047
: GENERAL INFORMATION:
: APPLICANT: Anderson, Burt E.
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
: TITLE OF INVENTION: DIAGNOSING
: TITLE OF INVENTION: ROCHALIMAEA HENSELAE AND ROCHALIMAEA QUINTANA
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NEEDLE & ROSENBERG, P.C.
: STREET: 127 Peachtree Street, Suite 1200
: CITY: Atlanta
: STATE: Georgia
: COUNTRY: USA
: ZIP: 30303
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: FILING DATE: US/08/245,294
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Spratt, Gwendolyn D.
: REGISTRATION NUMBER: 36,016
: REFERENCE/DOCKET NUMBER: 1414.612
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 404/688-9870
: TELEFAX: 404/688-9880
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1791 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 141..1649
US-08-245-294-7

Query Match 11.1%; Score 154.2; DB 1; Length 1791;
Best Local Similarity 50.4%; Pred. No. 2.1e-33;
Matches 403; Conservative 0; Mismatches 393; Indels 3; Gaps 1;

Qy 373 ttccggttcggcgttcacacgcaaaacggctacatcctgaccaaataaccacgctcgtt 432
Db 516 TTTGATCGGGTTTTTATCTCGTCTGATGTTATATTTGTGACCAATATCATGTGATT 575

Qy 433 gccggtatgggcagatcaaaagtcctgcctcaacgacgacgcaatataaccgccaaactc 492
Db 576 TCTGATGCCACAAGTTACGCTGTTCTTGTGATGACGGTACAGACTGATGCAAAACTC 635

Qy 493 atcgggttcggatgtccaatccgatcgcgccctcttgaataatgaacgcaacgaaagacta 552
Db 636 ATTGGACGGACCCACGAACTGATCTGCAGTATTAAAGTCAATGAAAGAAAAAATTT 695

Qy 553 cccgtcgtcaaatcggcaatcccccaaatgtgaacacggcggaatggctgcgcac 612
Db 696 TCGTACGTGATTTTGGTGATGATTCAAACATTCGTGTGGTATGGGTGTTGTGATT 755

Qy 613 ggcgcgccttcggttttgacaacacgctgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 672
Db 756 GGTATCATTTGGTCTTGGTGGAACCTGTGACAGCAGGTATCGTTTCACGACGTGGACGT 815
```

```
Qy 673 agcccgcccaacgaaagcctacacacccctlcalccaaacacgacgltlgccalcaatccgggc 732
Db 816 GATATCGGTACCGGTGTTTATGATGATGATTTTATTCAGATTGCTGCCAGTTAATCAGGA 875

Qy 733 aattccggcgcccgctgttcaacttaaaagacagctgcctcgccatcaattcgcgaata 792
Db 876 AATTCTGGAGGTCCAACTTTTGATCTTAACGGAAGGTTGTTGGAGTGAATACGGCAATT 935

Qy 793 tacagccgacgcgcggttcatctgggcattcctcttgcctccctccgattgacgttgcctg 852
Db 936 TTTTCTCTTCTGCGGCCAACGTTGGGATTGCTTTCGCTATTTCGCCGACCAACACGGAAC 995

Qy 853 aatgctgcgaacagctgaaaaacacccggcgaagLccaaacgcggacaactggcgtgatt 912
Db 996 GAGTTTGTGCAACAACATTATCGAAAAAGCTTTAGTTTACGCGTGGTTGGCTTGGGTTCAG 1055

Qy 913 attcaggaagtatcctacggtttggcacagctcgttcctgctctggtatgataaagccagcgca 972
Db 1056 ATTACAGCTGTAAACAAAGAAATTTCTGATTCAATTTGTTTGAAGGAGGCTAAAGGTGGG 1115

Qy 973 ttgattgccaaaatccttcccgcgagcccgccagacgctgcggcctgcaggcgagcgac 1032
Db 1116 TTAATTACCGATCCATTAAAGGGGCCACCCCAAAA---GCTGGTATCAAGGCACGTGAT 1172

Qy 1033 atcgtcctcagctcgcgcggcgagaaatacgtlcttccgcgcgaccttcccgltc 1092
Db 1173 GTTATTATTTCGGTAAATGCTGAGAAGATTAAATGATGTCCTGATCTAGCAAAAGCGTATT 1232

Qy 1093 ggcgcattacgcgggaaagaaagagtcagcctcgcgcgtatggcgcaaaagcgcaaatc 1152
Db 1233 GCAAAATATGAGCCCGAGGAGAAACAGTAACCTTAGGAGTGTGAAATCTGCTAAAGAGAG 1292

Qy 1153 acaatcaaacgcaagctgg 1171
Db 1293 AATATTAAAGTTAACTTG 1311

RESULT 5
US-08-474-499-7
: Sequence 7, Application US/08474499
: Patent No. 5693776
: GENERAL INFORMATION:
: APPLICANT: Anderson, Burt E.
: APPLICANT: Regnery, Russell L.
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
: TITLE OF INVENTION: DIAGNOSING ROCHALIMAEA HENSELAE
: TITLE OF INVENTION: AND ROCHALIMAEA QUINTANA INFECTION
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NEEDLE & ROSENBERG, P.C.
: STREET: 127 Peachtree Street, Suite 1200
: CITY: Atlanta
: STATE: Georgia
: COUNTRY: USA
: ZIP: 30303
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/474,499
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/245,294
: FILING DATE: 18-MAY-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Spratt, Gwendolyn D.
: REGISTRATION NUMBER: 36,016
: REFERENCE/DOCKET NUMBER: 1414.612
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 404/688-0770
```

```
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1791 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 141..1649
;
US-08-474-499-7

      Query Match      11.1%; Score 154.2; DB 1; Length 1791;
      Best Local Similarity 50.4%; Pred. No. 2.le-33;
      Matches 403; Conservative 0; Mismatches 393; Indels 3; Gaps 1;

QY 373 ttccgttcgggttcacatcagcaaaacgggtacatcctcgaccataccccacgtcgtt 432
Db 516 TTTGGATCGGGTTTTTTTATCTCGTCTGATGGTTATATTGTGACCAATAATCATGTGATT 575

QY 433 gccgggtatggcgagatcaaaagtcctgctcaacgacaaagcgcggaataataccgccaactc 492
Db 576 TCTGATGSCACAAGTTACGCTGTGTCTTCTGTATGACGGTACAGAACTGAATCCTGCTATT 635

QY 493 atcggttcgagatgccaatcgatcgcccttcgaaatcgcacgcaacggaagagcta 552
Db 636 ATTGGACGGACCCACGAACGTACTTGCAGTATTAAAAGTCAATGAAAAAGAAAAATTT 695

QY 553 cccgctcgtaaaatcggccaatcccccaaaatttgaacacggcggaatgggtcgtccatc 612
Db 696 TCGTACGTTGATTTTGTGTGATGATTCAAACTTCGCTGCTGTGATGGTGTGCTGATT 755

QY 613 ggcgcgccttcggctttgcaacacagcgtgaccccgccgacgtgctgcgcgaagagcaga 672
Db 756 GGTAAATCCATTGGTCTTTGGTGAACGTGTGACACAGGTATCGTTTTGACGACGTGGACGT 815

QY 673 agcctgcacacgaagctcacacaccccttcacaaacccgacgttgccatcaatccgggc 732
Db 816 GATATCGGTACCGGTGTTTATGATGATTTATTCAGATTGATGCTGCAGTTAATCGAGGA 875

QY 733 aattccggcgccgcgtgttcaacttaaaaggacagtcgctgcgcataatctcgaata 792
Db 876 AATTCTGGAGGTCCAACCTTTTGATCTTAACGGAAGGTTGTGGAGTGAATACGGCAATT 935

QY 793 tacagccgacgagcgagatcattgagcatctcctttgccatcccgattgacgttgccatg 852
Db 936 TTTTCTCTTCTGGGGGCAACGTTGGGATTCGCTATTCGCGACGACGACGCGAAC 995

QY 853 aatgctcgcgaacagctgaataaacacacggcgaagtcacacgagcagacaaactggcgatg 912
Db 996 GAGGTGTGCAACAACATTATCGAAAAAGGTTTAGTTCAGCGTGGTTGCGTTGGGGTTTCA 1055

QY 913 attcagaagtatcctcaggttttggcacagtcgttcggtctggataaaagccagcgcgca 972
Db 1056 ATTCAGCCTGTGAACAAAAAATTTCTGATTCAATTGGTTTGAAGAGGCTAAAAGTGGC 1115

QY 973 ttgattgccaaaatcccttcggcagccccgcagaaacgtgcgccttcagcgcgcgac 1032
Db 1116 TTAATTACCGATCCATTAAAGGGGCCACGCCGCAAAA---GCTGGGTATCAAGCGGATGAT 1172

QY 1033 atcgtccctcagccttcgacggcggaataacgttcttcggcgacccctcccgatggtc 1092
Db 1173 GTTATTATTCGGTAAATGGGTGAGNAGATTAAATGATGCTCCGTGATCTAGCAAGCGTATT 1232

QY 1093 ggcgcattacccggggaaaaagagtcagcctcgccgctatgctgcgaagcggaagaaatc 1152
Db 1233 GCMAATATGAGCCAGGAGAACAGTAACCTTAGGAGTTTGGAAATCTGGTAAAGAGAG 1292

QY 1153 acaatcaaaagcgaactgg 1171
Db 1293 AATATTAAAGGTTAAACTTTG 1311
```

```
RESULT 6
US-08-307-279A-7
; Sequence 7, Application US/08307279A
; Patent No. 5736347
; GENERAL INFORMATION:
;   APPLICANT: Anderson, Burt E.
;   APPLICANT: Regnery, Russell L
;   TITLE OF INVENTION: Nucleic Acids of Rochallimaea Henselae
;   TITLE OF INVENTION: and Methods and Compositions for Diagnosing Rochallimaea
;   TITLE OF INVENTION: Henselae and Rochallimaea Quintana Infection
;   NUMBER OF SEQUENCES: 14
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: NEEDLE & ROSENBERG, P.C.
;     STREET: 127 Peachtree Street, N.E., Suite 1200
;     CITY: Atlanta
;     STATE: Georgia
;     COUNTRY: USA
;     ZIP: 30303
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/307,279A
;     FILING DATE:
;     CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Spratt, Gwendolyn D.
;     REGISTRATION NUMBER: 36,016
;     REFERENCE/DOCKET NUMBER: 1414.624
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (404) 688-0770
;     TELEFAX: (404) 688-9880
;     INFORMATION FOR SEQ ID NO: 7:
;       SEQUENCE CHARACTERISTICS:
;         TYPE: nucleic acid
;         STRANDEDNESS: double
;         TOPOLOGY: linear
;         MOLECULE TYPE: DNA (genomic)
;         FEATURE:
;           NAME/KEY: CDS
;           LOCATION: 141..1652
;
US-08-307-279A-7
```

```
Query Match      11.1%; Score 154.2; DB 1; Length 1791;
Best Local Similarity 50.4%; Pred. No. 2.le-33;
Matches 403; Conservative 0; Mismatches 393; Indels 3; Gaps 1;

QY 373 ttccgttcgggttcacatcagcaaaacgggtacatcctcgaccataccccacgtcgtt 432
Db 516 TTTGGATCGGGTTTTTTTATCTCGTCTGATGGTTATATTGTGACCAATAATCATGTGATT 575

QY 433 gccgggtatggcgagatcaaaagtcctgctcaacgacaaagcgcggaataataccgccaactc 492
Db 576 TCTGATGSCACAAGTTACGCTGTGTCTTCTGTATGACGGTACAGAACTGAATCCTGCTATT 635

QY 493 atcggttcgagatgccaatcgatcgcccttcgaaatcgcacgcaacggaagagcta 552
Db 636 ATTGGACGGACCCACGAACGTACTTGCAGTATTAAAAGTCAATGAAAAAGAAAAATTT 695

QY 553 cccgctcgtaaaatcggccaatcccccaaaatttgaacacggcggaatgggtcgtccatc 612
Db 696 TCGTACGTTGATTTTGTGTGATGATTCAAACTTCGCTGCTGTGATGGTGTGCTGATT 755

QY 613 ggcgcgccttcggctttgcaacacagcgtgaccccgccgacgtgctgcgcgaagagcaga 672
Db 756 GGTAAATCCATTGGTCTTTGGTGAACGTGTGACACAGGTATCGTTTTCAGCAGCTGGACGT 815
```


Qy 673 agcctgccccagcgaagctacacacctctacccaaacccagcagtggtccatcaatcccgccg 732
Db 816 GATATCGGTACCGGTGTTTATGATGATTTTATTCAGATTGATGCTGCAGTTAATCGAGGA 875
Qy 733 aattccggcgcccgctgttcaactaaagacagcagtgctggcgaatcaattcgcaata 792
Db 876 AATTCTGGAGGTCCAACTTTTGTATCTTAACGGAAGGTTGTTGGAGTGAATACGGCAATT 935
Qy 793 tacagccgagcgcggtatctatgggcatctcttggccatcccgattgacgttgcctg 852
Db 936 TTTTCTCTCTCTGCGGGCAACGTTGGGATTTGCTTTCCGCTATTTCCGCGCACCAACAGCGAAC 995
Qy 853 aatgtcgccgaacagctgaacaaacccgcaaaagtccaacgcggaacaactggcggtgatt 912
Db 996 GAGGTCTGTCACAACTTATCAGAAAGGTTTGTATTCAGCGTGGTTGGGTTTCAG 1055
Qy 913 attcaggaagtatcctaggtttggcacagtgcttggctcgttgatgataaaagccagcgcca 972
Db 1056 ATTACGCTGTAAACAAAAGAAATTTCTGATTCAAATTTGGTTTGAAGGAGGCTAAAGGTGCG 1115
Qy 973 ttgattgccccaaatctctcccgccagcccgccgagaaacgtgcccggcctgagcgcgccgac 1032
Db 1116 TTAATTAACGATCCATTAAGGGGCGCCAGCGCAAAA--GCTGGTATCAAGGCGAGGTGAT 1172
Qy 1033 atcgtccctcagcctcgacggcggaagaaatacgttcttccggcgaccttcccgctcagtc 1092
Db 1173 GTTATTATTTCGGTAAATGGTGAGAGATTAATGATGTCGGTATCAGCAAAAGCGTATT 1232
Qy 1093 ggcgcattacggcggaagaaagcagctcggcgatgagcgaagcggaagaaatc 1152
Db 1233 GCAATATGAGCCCGAGGAACAGTAACCTTAGGAGTTTGGAAATCTGGTAAAGAAGAG 1292
Qy 1153 acaatcaaaagcgaagctgg 1171
Db 1293 AATATTAAAGTTAAACTTG 1311

RESULT 7

PCT-US95-06211-7
; Sequence 7, Application PC/TUS9506211
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: ROCHALIMAEA HENSELAE AND ROCHALIMAEA QUINTANA INFECTION
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESS: NEEDLE & ROSENBERG, P.C.
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06211
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,294
; FILING DATE: 18 MAY 1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwendolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 1414.6121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1791 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 141..1649
PCT-US95-06211-7

Query Match 11.1%; Score 154.2; DB 4; Length 1791;
Best Local Similarity 50.4%; Pred. No. 2.le-33;
Matches 403; Conservative 0; Mismatches 393; Indels 3; Gaps 1;

Qy 373 ttcgggttcgggcttcacatcagcaaaaacggctacatcctgaccaaataccacacutcggt 432
Db 516 TTTTGGATCGGGTTTTTTTATCTCGTCTGATGGTTTATTTGTGTCACCAATAATCATGTGATTT 575
Qy 433 gccgggtatgggcaglatcaaaagtcctgctcaacgcaagcgcaatataaccgccaactc 492
Db 576 TCTCATGGCACAAGTTACGCTGTGTCTTGTATGACGGTTACAGAACTGAATGCAAAACATC 635
Qy 493 atcgggttcggatgtcccaatccgagtgctgccccttctgaaatacgcgcgaagaaagcta 552
Db 636 ATTGGAACGGACCCACGAACTGATCTTCAGCTATTTAAAGTCAATGAAAGAAAGAAATTT 695
Qy 553 cccgtcgtcaaaaatcgccaatccccaaaatttgaaacccggcgcaatgggtcgctgcatc 612
Db 696 TCGTACGTTGATTTTGGTGATGATTCAAAACCTTCGTTTGGTGATTTGGTGTGCTATT 755
Qy 613 ggcgcgccttcggcttttgacaacagcgtaccgcgcgcgcgcgcgcgcgcgcgcgcgcgc 672
Db 756 GGTAATCCATTTGGTCTTGGTGGAACTGTACAGCAGGTATCGTTTTCAGCAGCTGCACGT 815
Qy 673 agcctgccccagcaagctacacacctctccaaacccgacgttgccatcaatccgggc 732
Db 816 GATATCGGTACCGGTGTTTATGATGATTTTATTCAGATTGATGCTGCAGTTAATCGAGGA 875
Qy 733 aattccggcgcccgctgttcaacttaaaagacaggtcgctcgccatcaatccgcaata 792
Db 876 AATTTCTGAGGTCCAACTTTTGTATTCAGGAAAGGTTTGTGAGTGAATACGCCAATT 935
Qy 793 tacagccgagcgcggtatcattcagggcctcctcttgcacatcccgatlgacgtlccatg 852
Db 936 TTTTCTCTCTCTGCGGGCAACGTTGGGATTTGCTTTCCGCTATTTCCGCGCACCAACAGCGAAC 995
Qy 853 aatgtcgccgaacagctgaacaaacccgcaaaagtcgaacgcgcgcgcgcgcgcgcgcgcgc 912
Db 996 GAGGTGTGCANCAACTTATTCGAAAGGTTTATGTCAGCGTGGTTGGCTTGGGGTTTCAG 1055
Qy 913 attcaggaagtatcctcagctgttgccacagctgttcggtctggaataaagccagcgcgca 972
Db 1056 ATTACGCTGTAAACAAAAGAAATTTCTGATTCAAATTTGTTTCAAGGAGGCTAAAGGTGCG 1115
Qy 973 ttgattgccccaaatccttcccgcgagcccccgcagcaactgcccgcgcgcgcgcgcgcgcgc 1032
Db 1116 TTAATTAACGATCCATTTAAAGGGCGCCAGCGCAAAA--GCTGGTATTCAGGAGGTGAT 1172
Qy 1033 atcgtcctcagcctcgacgc 1092
Db 1173 GTTATTATTTCGGTAAATGGTGAGNAGATTAATGATGTCGGTGTATCTAGCAAAAGCGTATT 1232
Qy 1093 ggcgcattacggcggaagaaagcagctcggcgatgagcgaagcggaagaaatc 1152
Db 1233 GCAAAATATGAGCCCGAGGAACAGTAACCTTAGCAGTTTGGAAATCTCGTAAAGAAGAG 1292
Qy 1153 acaatcaaaagcgaagctgg 1171
Db 1293 AATATTAAAGGTAAACTTG 1311

RESULT 8
US-08-278-091-1

```

; Sequence 1, Application US/08278091
; Patent No. 5506139
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin7 Protein with
; TITLE OF INVENTION: Reduced Protease Activity
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/278,091
; FILING DATE: 21-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-371
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-278-091-1

Query Match      10.0%; Score 139.4; DB 1; Length 2894;
Best Local Similarity 49.8%; Pred. No. 3.1e-29;
Matches 384; Conservative 0; Mismatches 381; Indels 6; Gaps

QY 382 ggcttcacatcagcaaaaaggctacatcctgacacaataaccacgctctgttcggtatg 441
Db 1108 GTCATTATTAATGCAAGCAAAAGGCTATGTTTAAACCAATAATCATGTTATTGATGAAGCT 1167
QY 442 ggcagatcaaaagtctgctcaacgacgaagcgcaatataccgcgaacctcatcggttcg 501
Db 1168 GATAAATTACCGTCANTTACAAATGCGGCTGTGATTTAAAGCAAAATAGTGGGTAA 1227
QY 502 gatgccaatccgatgtgcgcctctgaaaatcgacgcgaacggaagagctaccgcgtcgc 561
Db 1228 GATCAACATATCAGATATTGCATTAGTACAGCTTGAAAAACCAAGTAATTTAACAGAAATC 1287
QY 562 aaatcgggcaatcccaaaaatttgaaacccggcgcaatgggtcgctgccatcgcgccgc 621
Db 1288 AAATTTGCTGATTCCGCAGAAATTAGCGGTAGCGGATTTTCACTGTTTCAATCGGTAATCCA 1347
QY 622 ttcggtcttgacaacagcgtagcccgcgcatcgctgctcgcccaagg-----cagaagc 675
Db 1348 TTTGGTTTAGGTCAACCTTGATCAGCATGAGGTATGTTTCTGCAFTGGGTGCTTCAACAGGT 1407
QY 676 ctgcccacgaaagctacacaccccttcacccaaacgacggttgccatcaatccgggcaat 735
Db 1408 TCTGACAGTGGCACTTATCAAAACTATATTCAACCGGATGCAGCAGTAACCCGGGTAA 1467
QY 736 tccggcgcccgctgttcaacttaaaagacaggtgctcgggcatcaatctcgcaaatatc 795

```

STRANDEDNESS: single
TOPOLOGY: linear
US-08-483-859-1

Query Match 10.0%; Score 139.4; DB 1; Length 2894;
Best Local Similarity 49.8%; Pred. No. 3.1e-29;
Matches 384; Conservative 0; Mismatches 381; Indels 6; Gaps 1;

```

Qy 382 ggcttcacacagcaaaacggctacatccctgacacaaataccacacgctgtgcccgtatg 441
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1108 GTCATTATTAATGCAAGCAAGGCTATGTTTAAACCAATAATCATGTTATTGATGAAGCT 1167

Qy 442 ggcaglatcaaaagtcctgctcaacgacagcgcaatataccgcccacaaactcatcggtcg 501
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1168 GATAAAATTAACGTCGAATTTACAAGATGGCGCTGAATTTAAAGCAAAATTTAGTGGGTAAA 1227

Qy 502 gatgtccaatccgatgtcgcccttctgaaaaatcgacgcaacggaagagctaccgcgtc 561
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1228 GATGAACATATCAGATATTTGCAATTAGTACAGCTTGAAAAACCAAGTAATTTAAACAGAAATC 1287

Qy 562 aaaaatcggaatcccaaaaattgaaacggcggaatgggtcgctgcccacatcgcgcgcc 621
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1288 AAATTTGCTGATTCGACAAATTAGCGGTAGCGGATTTTCACTGTTGCAATCGGTATCCA 1347

Qy 622 ttcggctttgacaaacagcgctgaccgcccgcacgtcggtgtccgccaagg-----cagaagc 675
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1348 TTTGGTTTAGTTCAAACTGTGACATCAGGTATGTTTCTGCAATTTGGTGTTCACACAGGT 1407

Qy 676 ctgcccacgaagactcacacccttccaaacggcggaatgggtcgctgcccacatcgcaaatatc 735
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1408 TCTGACATGGCATTATGAAAACCTATATTTCAAAACGATGTCAGCAGTAAACCCGCGTAAT 1467

Qy 736 tccggcgccgctgttcaaaactaaagacaggtcgctcgccacatcgcaaatatcgcaaatatc 795
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1468 TCGGTGAGCGTTAGTAACTTAATGCGCACTTATTTGGAATTAATACCGCAATATT 1527

Qy 796 agccgacgagcggtattcatgggcatctcttggccatccccgattgacgttgccatgaat 855
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1528 TCTCAAGCGGTGGCAATGCAAGAAATGCTTTGGGATTTCCAAGTAATCAAGCAAGCAAT 1587

Qy 856 gtcgccgaagactgaaacacacggcgaagatcccaacgagcaactggcggtgattatt 915
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1588 TTAGTGCAACAAATTTTGAATTTGGTCAAGTGGTGGGATTTGCTTGGTATTAAAGGT 1647

Qy 916 caggaagtactcactcggttggcacagctgcttgcgtgctgataaaagccgagcgcgcatg 975
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1648 GCGCAACTCAATGCTGATTTAGCCAAAGCCTTTAATGTAAGCGCGCAACAGGCGCATTT 1707

Qy 976 attgccccaaatccttcccgccagcccccgacagacgtgcccgtcgagggcgagacalc 1035
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1708 GTAAGTGAAGTTTACCAGAAATCTGCTGCTGAAAAAGCAGGACTTTAAAGCGGCGGATATT 1767

Qy 1036 gtccctcagcctcgacgagcaggaatacgttcttccggcgaccttccctcatggtcgcc 1095
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1768 ATCAGGCGGATGAACGGTCAAAAATCTCAAGTTTCGCTGAAATTCGTGCAAAAATCGCA 1827

Qy 1096 gccattacgcccgggaaaaagaagtcagcctcgcgctatggcgcaaaagcgaa 1146
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1828 ACCACTGTGCAAGGCAAGAGATTAGCTTGACTTACTTACTTACTTACTTACTTACTTACTT 1878

```

RESULT 10

US-08-472-173-1
Sequence 1, Application US/08472173
Patent No. 5665353
GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M
APPLICANT: YANG, Yan-Ping
APPLICANT: CHONG, Pele
APPLICANT: OOMEN, Raymond P.
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with

```

; TITLE OF INVENTION: Reduced Protease Activity
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472.173
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296.149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278.091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-493 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-472-173-1

```

Query Match 10.0%; Score 139.4; DB 1; Length 2894;
Best Local Similarity 49.8%; Pred. No. 3.1e-29;
Matches 384; Conservative 0; Mismatches 381; Indels 6; Gaps 1;

```

Qy 382 ggcttcacacagcaaaacggctacatccctgacacaaataccacacgctgtgcccgtatg 441
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1108 GTCATTATTAATGCAAGCAAGGCTATGTTTAAACCAATAATCATGTTATTGATGAAGCT 1167

Qy 442 ggcaglatcaaaagtcctgctcaacgacagcgcaatataccgcccacaaactcatcggtcg 501
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1168 GATAAAATTAACGTCGAATTTACAAGATGGCGCTGAATTTAAAGCAAAATTTAGTGGGTAAA 1227

Qy 502 gatgtccaatccgatgtcgcccttctgaaaaatcgacgcaacggaagagctaccgcgtc 561
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1228 GATGAACATATCAGATATTTGCAATTAGTACAGCTTGAAAAACCAAGTAATTTAAACAGAAATC 1287

Qy 562 aaaaatcggaatcccaaaaattgaaacggcggaatgggtcgctgcccacatcgcgcgcc 621
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1288 AAATTTGCTGATTCGACAAATTAGCGGTAGCGGATTTTCACTGTTGCAATCGGTATCCA 1347

Qy 622 ttcggctttgacaaacagcgctgaccgcccgcacgtcggtgtccgccaagg-----cagaagc 675
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1348 TTTGGTTTAGTTCAAACTGTGACATCAGGTATGTTTCTGCAATTTGGTGTTCACACAGGT 1407

Qy 676 ctgcccacgaagactcacacccttccaaacggcggaatgggtcgctgcccacatcgcaaatatc 735
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1408 TCTGACATGGCATTATGAAAACCTATTTCAAAACGATGTCAGCAGTAAACCCGCGGTAAAT 1467

Qy 736 tccggcgccgctgttcaaaactaaagacaggtcgctcgccacatcgcaaatatcgcaaatatc 795
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1468 TCGGTGAGCGTTAGTAACTTAATGCGCACTTATTTGGAATTAATACCGCAATATT 1527

Qy 796 agccgacgagcggtattcatgggcatctcttggccatccccgattgacgttgccatgaat 855
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Query Match	10.0%	Score 139.4	DB 2	Length 2894
Best Local Similarity	49.8%	Pred. No. 3.1e-29		
Matches 384	Conservative 0	Mismatches 381	Indels 6	Gaps 1
Qy 382	ggcttcacatcagcaaaaacggtacatcccgaccacatcccacgtctgttcgcggtatg	441		
Db 1108	GTCATTTAATGCAAGCAAAAGCGTATGTTTAAACCAATAATCATGTTATTGATGAAGCT	1167		
Qy 442	ggcagtatcaaaagtccctgctcaacagcaacgacgcaaatataccgccaactcatcggtctg	501		
Db 1168	GATAAATATACCGTGCAAATACAGATGGCGGTGAATTTAAAGCAAAATAGTGGGTAAA	1227		
Qy 502	gatgtccaatccgatgctgcctctctgaataatcgacgcacacggaagagctaccgcgtgc	561		
Db 1228	GATGAACATCAGATATTGCATATTAGTACAGCTTGAAAAACCAACGATTAATTTAACAGAAATC	1287		
Qy 562	aaatcggcaatcccaaaaatttgaaccggcggaattgctgcgtccatcggcgcgccc	621		
Db 1288	AAATTTGCTGATTCGCACAAATACGGGTAGGCGATTTCACTGTTCGAATCGGTAATCCA	1347		
Qy 622	ttcggcttgacaacagcgtgacgcgcggcatcgctccgcgaag-----cagaagc	675		
Db 1348	TTTGTTTAGTCAAACTGTCAATCAGGTATGTTTCTGCTATTTGGTTCGTTCAACAGGT	1407		
Qy 676	ctgcccacgaaagtacacaccccttcatccaaaccgagcgttgccataatccgggcaat	735		
Db 1408	TCTGACAGTGCACACTATGAAACATATATTCAAACCGATGCAGCAGTAACCGCGGTAA	1467		
Qy 736	tccggcgccgcgtgttcaactttaaaggacaggtcgtcggcatacaattcgcaataatc	795		
Db 1468	TCGGGTGGAGCGTGTAGTAACCTTAATTTGGCGAACCTTATTGCAATTAATACCGCAATTT	1527		
Qy 796	agccgacgcgcggtatcataggcatctcttggccatcccgattgacgttgcgatgaat	855		
Db 1528	TCTCAAGCGGTGCGAATGCAAGGAATTTGCCTTTTCGATTTCCAAAGTAATCAACGAA	1587		
Qy 856	gtcgcgcgaacagctgaaaaaacacgcggcaaaagtccaaacgcgacactggcggtatt	915		
Db 1588	TTAGTGCAACAAATTTTAGAATTTGGTCAACTGCGTCGCGATTGCTTGGTATTAAGGT	1647		
Qy 916	caggaagtatcctacggtttggacagtcgtctcggtctcgataaagccagcgcgcatcg	975		
Db 1648	GGCGAACTCAATGCTGATTTAGCCAAAGCCCTTTAATGTAAAGCGCGCAACAAGGCGCAT	1707		

Qy	622	ttcggctttgcaacacagcgtgaccgcgcggcatcgtgtccgcgcaaaagg-----cagaagc	675
Db	1348	TTTGGTTTAGGTCAAACTGTGACATCAGGTATTGTTTTCGATTGGGTTCGTTCAACAGGT	1407
Qy	676	ctgcgcaacgaagctacacaccttcatacaaccacagcttgcatacaatccgggcaat	735
Db	1408	TCGTGACGTGSCACTTATGAACAATATATTCAACCGNATGCAGCAGTAACCCGGGTAA	1467
Qy	736	tcggggcgccgctgtttcaacttaaaaggacagctcgtcggcatacaattcgcgaatatcac	795
Db	1468	TCGGGTGAGCGGTAGTAACCTTAATGGCGAAGCTATTATGGANATTAATACCGCAATTATT	1527
Qy	796	agcgcagcgcgcgcatcattggcgatctctttgccatcccgattgcgttgcacataat	855
Db	1528	TCTCAACGGTGGCAATGCAAGTAATGCCCTTTGGGATTCCAAGTAATCAACGCAAGCAAT	1587
Qy	856	gtcgcgcaacagctgaaaaaacccgcgcaaaagtccaacgcgacactggcggtgattatt	915
Db	1588	TTAGTGCACAATAATTTAGAATTTGGTCAAGTCGCGTCGCGATTGCTTTGGTATTAAAGGT	1647
Qy	916	caggaagratcctacggtttggacacagtcggttcgggtctggataaagccagcggcgcatcg	975
Db	1648	GGCGAACTCAATGCTGATTATGACCAAAAGCCCTTTAATGTAAAGCGCGCAACAAGGCGCATTT	1707

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: March 22, 2001, 02:20:35 ; Search time 899.71 Seconds
(without alignments)
10865.107 Million cell updates/sec

Title: US-09-388-090-3
Perfect score: 1395
Sequence: 1 gtgttcaaaaaataccaata.....tcaggcgcaagtcacctgcaa 1395

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7991742 seqs, 3503743858 residues 15983484

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: em_estba:*
42: em_estfun:*
43: em_esthum1:*

44: em_esthum2:*
45: em_esthum3:*
46: em_esthum4:*
47: em_esthum5:*
48: em_esthum6:*
49: em_esthum7:*
50: em_esthum8:*
51: em_esthum9:*
52: em_esthum10:*
53: em_esthum11:*
54: em_esthum12:*
55: em_esthum13:*
56: em_esthum14:*
57: em_esthum15:*
58: em_esthum16:*
59: em_esthum17:*
60: em_esthum18:*
61: em_esthum19:*
62: em_esthum20:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estov1:*
68: em_estov2:*
69: em_estpl1:*
70: em_estpl2:*
71: em_estpl3:*
72: em_estpl4:*
73: em_estpl5:*
74: em_estro1:*
75: em_estro2:*
76: em_estro3:*
77: em_estro4:*
78: em_estro5:*
79: em_estro6:*
80: em_estro7:*
81: em_estro8:*
82: em_estro9:*
83: em_estro10:*
84: em_estro11:*
85: em_estro12:*
86: em_estro13:*
87: gb_est41:*
88: gb_est42:*
89: gb_est43:*
90: gb_est44:*
91: gb_est45:*
92: gb_est46:*
93: gb_est47:*
94: gb_est48:*
95: gb_est49:*
96: gb_est50:*
97: gb_est51:*
98: gb_est52:*
99: gb_est53:*
100: gb_est54:*
101: gb_est55:*
102: gb_est56:*
103: gb_est57:*
104: gb_est67:*
105: gb_est68:*
106: gb_est69:*
107: gb_est70:*
108: gb_est71:*
109: gb_est72:*
110: gb_est73:*
111: gb_est74:*
112: em_esthum21:*
113: em_esthum22:*
114: em_esthum23:*
115: em_estom1:*
116: em_estom2:*

117: em_estpl6:*
 118: em_estpl7:*
 119: em_estpl8:*
 120: em_estpl9:*
 121: em_estpl10:*
 122: em_estpl11:*
 123: em_estpl12:*
 124: em_estpl13:*
 125: em_estpl14:*
 126: em_estpl15:*
 127: em_estpl16:*
 128: em_estpl17:*
 129: em_estpl18:*
 130: em_estpl19:*
 131: em_estpl20:*
 132: em_estpl21:*
 133: em_estpl22:*
 134: em_estpl23:*
 135: em_estpl24:*
 136: em_estpl25:*
 137: em_estpl26:*
 138: em_estpl27:*
 139: em_estpl28:*
 140: em_estpl29:*
 141: em_estpl30:*
 142: em_estpl31:*
 143: em_estpl32:*
 144: em_estpl33:*
 145: em_estpl34:*
 146: em_estpl35:*
 147: em_estpl36:*
 148: em_estpl37:*
 149: em_estpl38:*
 150: em_estpl39:*
 151: em_estpl40:*
 152: em_estpl41:*
 153: em_estpl42:*
 154: em_estpl43:*
 155: em_estpl44:*
 156: em_estpl45:*
 157: em_estpl46:*
 158: em_estpl47:*
 159: em_estpl48:*
 160: em_estpl49:*
 161: em_estpl50:*
 162: em_estpl51:*
 163: em_estpl52:*
 164: em_estpl53:*
 165: em_estpl54:*
 166: em_estpl55:*
 167: em_estpl56:*
 168: em_estpl57:*
 169: em_estpl58:*
 170: em_estpl59:*
 171: em_estpl60:*
 172: em_estpl61:*
 173: em_estpl62:*
 174: em_estpl63:*
 175: em_estpl64:*
 176: em_estpl65:*
 177: em_estpl66:*
 178: em_estpl67:*
 179: em_estpl68:*
 180: em_estpl69:*
 181: em_estpl70:*
 182: em_estpl71:*
 183: em_estpl72:*
 184: em_estpl73:*
 185: em_estpl74:*
 186: em_estpl75:*
 187: em_estpl76:*
 188: em_estpl77:*
 189: em_estpl78:*

190: gb_gss25:*
 191: gb_gss26:*
 192: gb_gss27:*
 193: gb_gss28:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	100.6	7.2	490	8	AA479844	AA479844 zu43h07.r
2	98.6	7.1	479	91	AA491392	AA491392 UI-M-BH3-
3	97.4	7.0	489	2	AA110074	AA110074 mo51b12.r
4	97.4	7.0	660	94	AA742283	AA742283 up55a10.y
5	95.4	6.8	456	16	AI159029	AI159029 vz83a08.r
6	94.6	6.8	521	93	AW654015	AW654015 103226 MA
7	92.2	6.6	474	146	W17262	W17262 zb17a08.r
8	92.2	6.6	485	144	R52327	R52327 yg75a09.r
9	91.9	6.5	504	93	AW660844	AW660844 99933 MAR
10	89.6	6.4	350	140	F07134	F07134 HSC1WG081.n
11	89.6	6.4	529	37	AV614726	AV614726 AV614726
12	87.8	6.3	559	37	AV599976	AV599976 AV599976
13	86.2	6.2	337	26	AI903651	AI903651 OV-RT032-
14	84.2	6.0	518	144	R22703	R22703 y08e06.r
15	81.2	5.8	288	146	T80106	T80106 y04h09.r
16	81.2	5.8	346	139	CI8770	CI8770 C18770 Iluma
17	80.2	5.7	500	38	AV667427	AV667427 AV667427
18	78.8	5.6	375	10	AA644825	AA644825 vs83a09.r
19	78.8	5.6	689	14	AF179474	AF179474 AF179474
20	77.8	5.6	623	27	AI982113	AI982113 PAT-PK007
21	77.2	5.5	280	89	AW326487	AW326487 19115 MAR
22	75.8	5.4	573	105	BE236269	BE236269 143992 MA
23	75.8	5.4	574	105	BE236299	BE236299 144028 MA
24	75.8	5.4	402	37	AV614725	AV614725 AV614725
25	72.4	5.2	509	90	AW62478	AW62478 BP230009B
26	69.6	5.0	292	6	AA349615	AA349615 EST56620
27	69.4	5.0	485	3	AA186646	AA186646 zp64g02.r
28	66.6	4.8	281	140	F08513	F08513 HSC3WA091.n
29	66.2	4.7	483	107	BE335671	BE335671 us87h06.y
30	65.6	4.7	508	141	H42958	H42958 yo68e01.r
31	65.4	4.7	496	38	AW013197	AW013197 Sp56f win
32	64.2	4.6	591	37	AV593667	AV593667 AV593667
33	64.2	4.6	544	134	BE026352	BE026352 db37g11.y
34	62.6	4.5	669	8	AA545554	AA545554 vj90d11.s
35	61.8	4.4	511	1	AA028600	AA028600 mil14a10.r
36	59.2	4.2	442	141	H16043	H16043 y127q02.r
37	57.4	4.1	504	13	AA913980	AA913980 vy91c06.r
38	57.4	4.1	275	5	AA340884	AA340884 EST46223
39	56.8	4.1	502	135	BE754831	BE754831 208470 MA
40	55.8	4.0	383	90	AW462230	AW462230 BP230008B
41	55.6	4.0	298	4	AA254878	AA254878 mz75h12.r
42	54.4	3.9	341	136	BE843509	BE843509 CM3-TN006
43	53.2	3.8	617	134	BE055938	BE055938 945020D03
44	52.2	3.7	427	93	AW632119	AW632119 91674 MAR
45	52.2	3.7	453	90	AW446594	AW446594 86442 MAR

ALIGNMENTS

RESULT 1
 AA479844
 LOCUS zu43h07.r1 Soares ovary tumor mRNA EST 09-NOV-1997
 DEFINITION IMAGE:740797 5' similar to SW:HTRA_SALTY P26982 PROTEASE DO
 PRECURSOR : mRNA sequence.
 AA479844
 ACCESSION AA479844.1 GI:2205730
 VERSION
 KEYWORDS EST.

converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH_BMAP_M_S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996).

TAG_LIB-NIH_BMAP_M_S4

TAG_TISSUE=corpus-striatum

TAG_SEQ=ACGGC

BASE COUNT 99 a 126 c 111 g 143 t
ORIGIN

Query Match 7.1%; Score 98.6; DB 91; Length 479;
Best Local Similarity 54.4%; Pred. No. 2.5e-17;
Matches 228; Conservative 0; Mismatches 179; Indels 12; Gaps 1;

```

QY 376 ggttcgggttcacatcagcaaaacggctacatcctctgaccaaataccacacgctgtgc 435
    || || || || || || || || || || || || || || || || || || || || || ||
Db 475 GGTTCAGGATTTCATCGTATCGGAGGATGGACTGATTGTGACAAATGCTCAGGTGGTCACC 416
    || || || || || || || || || || || || || || || || || || || || || ||
QY 436 ggtatgggcagttatcaaaagtctgtcgaacgacgaagcggaatataccgccaacatcatc 495
    || || || || || || || || || || || || || || || || || || || || || ||
Db 415 AACAAAAACCGGGTCAAGTTTCAGCTGAAGATGGAGTACCTATCAAGCCAAAAATCAAG 356
    || || || || || || || || || || || || || || || || || || || || || ||
QY 496 ggttcggatgtccaatccgatgtgcctcttctgaaaaatcgacgcaacggaagactacc 555
    || || || || || || || || || || || || || || || || || || || || || ||
Db 355 GATGTGGATGAAAGCGCGGACATTCGGCTTATCAAGATTGACCAACGAGGAAGCTGCCA 296
    || || || || || || || || || || || || || || || || || || || || || ||
QY 556 gtctcaaaatcggcaatcccaaaattgaaacggcggaatgggtgcctccatcgcc 615
    || || || || || || || || || || || || || || || || || || || || || ||
Db 295 GTCTGTCTGTCGGCGGCTCTCAGAGCTGAGACCTGGAGAAATTTGTAGTTGCCATTTGA 236
    || || || || || || || || || || || || || || || || || || || || || ||
QY 616 gcgcctctcggtttgacaacagcgtgacgcggtgcctctgtccgccaagagcagaagc 675
    || || || || || || || || || || || || || || || || || || || || || ||
Db 235 AGCCCTTTCTTCTTCAAAACACACTCACCTACCTGGGATCGTCAGCACCCACCGCAGGC 176
    || || || || || || || || || || || || || || || || || || || || || ||
QY 676 -----ctgccaacgaaagctacacaccccttcacaaacgacagcttgcctc 723
    || || || || || || || || || || || || || || || || || || || || || ||
Db 175 GGCAAGAGCTGGGACTTCGGAACCTCCGATATGGACTACATTACAGACAGACGCTATCATC 116
    || || || || || || || || || || || || || || || || || || || || || ||
QY 724 aatcgggcaattccgcggccgcgtgttcaacttaaaaggacaggtcgtcgccatcaa 782
    || || || || || || || || || || || || || || || || || || || || || ||
Db 115 ANTATGGAATTCGGAGGCCCGTTAGTAACCTGGATGGCGAGGTGATGGGATTAA 57
    || || || || || || || || || || || || || || || || || || || || || ||

```

RESULT 3

AAL10074 489 bp mRNA EST 06-NOV-1996
LOCUS m051b12.r1 Life Tech mouse embryo 10 5dpc 10665016 Mus musculus
DEFINITION cDNA clone IMAGE:557087 5' similar to SW:HTRA_SALTY P26982 PROTEASE
DO PRECURSOR ;, mRNA sequence.

ACCESSION AAL10074

VERSION AAL10074.1 GI:1661822

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 489)

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The WashU-HHMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:337879

Seq primer: -28M13 rev1 from Amersham

High quality sequence stop: 446.

FEATURES
Source

Location/Qualifiers
1..489
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:557087"
/clone_lib="Life Tech mouse embryo 10 5dpc 10665016"
/tissue_type="embryo"
/dev_stage="10.5dpc embryos"
/lab_host="DH10B"
/note="Organ: whole embryo; Vector: pCMV-SPORT2; Site_1:
Sall; Site_2: NotI; Cloned unidirectionally. Primer:
Oligo dt. 10.5dpc embryos. pCMV-SPORT2 vector."
BASE COUNT 130 a 116 c 140 g 103 t
ORIGIN

Query Match 7.0%; Score 97.4; DB 2; Length 489;
Best Local Similarity 53.8%; Pred. No. 5.5e-17;
Matches 231; Conservative 0; Mismatches 186; Indels 12; Gaps 1;

```

QY 366 attgaacttcggttcggttcacatcagcaaaacggctacatcctctgaccaaataccca 425
    || || || || || || || || || || || || || || || || || || || || || ||
Db 46 AGTGGCCAGTGGTTCAGGATTCATCGTATCGGAGGATGGAGTGTGTGACAAATGCTCA 105
    || || || || || || || || || || || || || || || || || || || || || ||
QY 426 cgtcgttcggttcggttcggttcggttcggttcggttcggttcggttcggttcggttc 485
    || || || || || || || || || || || || || || || || || || || || || ||
Db 106 CGTGGTCAACCAACAAACCGGGTCAAGGTTGAGCTGAAGAAATGGAGCTACCTATGAAGC 165
    || || || || || || || || || || || || || || || || || || || || || ||
QY 486 caaacatcatcggttcggttcggttcggttcggttcggttcggttcggttcggttcggttc 545
    || || || || || || || || || || || || || || || || || || || || || ||
Db 166 CAAATCAAGGATGTGGATGAAAGCGGACATTCGGGCTTATCAAGATTGACCAACGAGG 225
    || || || || || || || || || || || || || || || || || || || || || ||
QY 546 agagctaccctgcgtcgaatccgcaatcccaaaattgaaacggcggttcggttcggttcg 605
    || || || || || || || || || || || || || || || || || || || || || ||
Db 226 AAAGTCCAGCTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 285
    || || || || || || || || || || || || || || || || || || || || || ||
QY 606 tgcacatcggttcggttcggttcggttcggttcggttcggttcggttcggttcggttcggttc 665
    || || || || || || || || || || || || || || || || || || || || || ||
Db 286 TGCCATTGGAAAGGCCCTTTCTCTTCAAAACACACTCACCTACCTGCGTACGACACCA 345
    || || || || || || || || || || || || || || || || || || || || || ||
QY 666 agcaaaagc-----ctgccaacgaaagctacacaccccttcacaaacgga 713
    || || || || || || || || || || || || || || || || || || || || || ||
Db 346 CCAGCGAGGCGGCAAAAGAGCTGGGACTTCGGAACCTCCGATATGGACTACATTACAGACAGA 405
    || || || || || || || || || || || || || || || || || || || || || ||
QY 714 cgttgccatcaatccggttcggttcggttcggttcggttcggttcggttcggttcggttcg 773
    || || || || || || || || || || || || || || || || || || || || || ||
Db 406 CGCTATCATCAATTATGGAGATTCCGGAGGCCCGTTAGTAAACCTGGATGGCGAGGTGAT 465
    || || || || || || || || || || || || || || || || || || || || || ||
QY 774 cggcatcaa 782
    || || || || || || || || || || || || || || || || || || || || || ||
Db 466 TGGGATTAA 474
    || || || || || || || || || || || || || || || || || || || || || ||

```

RESULT 4

AW742283

LOCUS

DEFINITION

sequence.

ACCESSION AW742283

VERSION AW742283.1

KEYWORDS

SOURCE

ORGANISM

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 660)

REFERENCE

AW742283 660 bp mRNA EST 07-SEP-2000
up55a10.y1 Soares_mouse NMIE Mus musculus cDNA clone IMAGE:2779962
5' similar to TR:092743 092743 NOVEL SERINE PROTEASE. ;, mRNA
sequence.

AW742283 660 bp mRNA EST 07-SEP-2000
up55a10.y1 Soares_mouse NMIE Mus musculus cDNA clone IMAGE:2779962
5' similar to TR:092743 092743 NOVEL SERINE PROTEASE. ;, mRNA
sequence.

AW742283 660 bp mRNA EST 07-SEP-2000
up55a10.y1 Soares_mouse NMIE Mus musculus cDNA clone IMAGE:2779962
5' similar to TR:092743 092743 NOVEL SERINE PROTEASE. ;, mRNA
sequence.

AW742283 660 bp mRNA EST 07-SEP-2000
up55a10.y1 Soares_mouse NMIE Mus musculus cDNA clone IMAGE:2779962
5' similar to TR:092743 092743 NOVEL SERINE PROTEASE. ;, mRNA
sequence.

AW742283 660 bp mRNA EST 07-SEP-2000
up55a10.y1 Soares_mouse NMIE Mus musculus cDNA clone IMAGE:2779962
5' similar to TR:092743 092743 NOVEL SERINE PROTEASE. ;, mRNA
sequence.

AW742283 660 bp mRNA EST 07-SEP-2000
up55a10.y1 Soares_mouse NMIE Mus musculus cDNA clone IMAGE:2779962
5' similar to TR:092743 092743 NOVEL SERINE PROTEASE. ;, mRNA
sequence.

Qy	496	ggttcgatgtccaatcogatgtctgcgccctctgaaatcgcagcaacgaagagctacccc	555
Db	130		
Db	130	GATGTGATGAAAGGGGGGACATTGCGCTTATCAAGATTGACCACGAGGAAAGCTGCCA	189
Qy	556	gtcgtcaaaatcggcaatccccaaatttgaaccggcggaatagggtcgtcgcacatgcgc	615
Db	190		
Db	190	GTCCTGCTGCTCGGCCGCTCTCAGAGCTGAGACTGGAGAAATTTGTATGTTGCCATTGGA	249
Qy	616	ggcctctcggtcttgcacaacagcgtgaccgcgcggcatcgtgtccgcgaacgaagcagaagc	675
Db	250		
Db	250	AGCCCTTTCTCTTCANACACAGTACCACCTGGGATCGTCAGCACCAACCCAGCGAGGC	309
Qy	676	-----ctgcccacgaagctacacaccccttcacccaaaccgaagcttgcacatc	723
Db	310	GGCAAGAGCTGGGACTTCGGAATCCGGATATGGACTACATTCAGACAGACGCTATCATC	369
Qy	724	aatccgggcaattccggcgccgcctgttcaacttaaaaggacagaggtcgtcgggaatcaa	782
Db	370		
Db	370	TAATTATGAAATTCGGAGCGCCGTTAGTAAACCTGGATGCGGAGGTGATTTGGGATTAA	428
RESULT 6			
AW654015 521 bp mRNA EST 14-JUL-2000			
LOCUS 103226 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.			
DEFINITION AW654015			
ACCESSION AW654015			
VERSION AW654015.1 GI:7419841			
KEYWORDS EST.			
SOURCE cow.			
ORGANISM Bos taurus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos. 1 (bases 1 to 521) Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid ,W.W. and Keele,J.W. Design and use of four pooled tissue normalized cDNA libraries for EST discovery in cattle Unpublished (2000) Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options. PCR Primers FORWARD: AGGAACACGCTATGACCAT BACKWARD: GTTTCGCCAGTCACGAG Plate: 103 row: D column: 5 Seq primer: ATTTAGGTGACACTATAG. Location/Qualifiers 1. 521 /organism="Bos taurus" /db_xref="taxon:9913" /clone_lib="MARC 1BOV" /tissue_type="pooled" /lab_host="DH10B" /note="vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI; Library made from pooled tissue from lymph node, ovary, fat, hypothalamus, and pituitary." BASE COUNT 134 a 135 c 147 g 105 t ORIGIN			
Query Match 6.8% Score 94.6; DB 93; Length 521;			
Best Local Similarity 53.1%; Pred. No. 3.5e-16;			
Matches 233; Conservative 0; Mismatches 194; Indels 12; Gaps 1;			
Qy	366	atgaacttcggttcgggttcattcatcagcaaaacggctacatccctgcaccaataccaca	425

strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGGAGCGCGCAATTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p773 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. This library was constructed
from the same fetus as the fetal heart library, Soares
fetal heart NBH19W." 93 t 3 others

BASE COUNT 122 a 121 c 135 g 93 t 3 others
ORIGIN

Query Match 6.6%; Score 92.2; DB 146; Length 474;
Best Local Similarity 57.4%; Pred. No. 1.6e-15;
Matches 166; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 376 ggttcgggttcacatcagcaaaacggctacatcctgacccaatacccaacgctcgttgc 435
DB 41 GGGTCGTGGGTTTATTGTCGGAAGATGGACTGATCGTGACAAATGCCACGTGGTGACC 100
QY 436 ggtatgggcagtcataaagtcctgtcacaacgacgaatataccgccaactcattc 495
DB 101 AACAGCAGCGGGTCAAGGTTGAGCTGAAGACGGTGCCACTTACGAAGGCCAAATCAAG 160
QY 496 ggttcggatgtccatcccatgctgcctctctgaaatcgacgcaacggaagagctacc 555
DB 161 GATGTGGATGAGAAACGACATCGGCACTCATCAAAATTGACCAACGAGGCAAGCTGCCT 220
QY 556 gtcgtcaaaatcgcaatcccaaaatttgaaacccgggggaatgggtcgctgcacatcgc 615
DB 221 GTCCTGCTGCTGGCGCTCTCAGAGCTGCGCGCGGAGAGTTGCTGTGTCGCCATCGGA 280
QY 616 ggcgccttcggctttgacaacagcgtgacgcgcgggcatcgtgtccgccca 664
DB 281 AGCCCGTTTTTCCCTTCAAAACACAGTCACACCGGGATCGTGAGCACCA 329

RESULT 8
R52327 485 bp mRNA EST 18-MAY-1995
LOCUS y975a09.r1 Soares infant brain lN1B Homo sapiens cDNA clone
DEFINITION IMAGE:38967 5' similar to SP:HTRA_SALTY P26982 PROTEASE DO
PRECUSOR , mRNA sequence.
ACCESSION R52327.1 GI:814229
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 485)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasaki,E., Waterston
R., Williamson,A., Wohlmann,P. and Willson,R.
TITLE The WashU-Werck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1476
High quality sequence stops: 391 Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1476 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 391.
Location/Qualifiers

source

1. 485
/organism="Homo sapiens"
/db_xref="GDB:411508"
/db_xref="taxon:9606"
/clone="IMAGE:38967"
/clone_lib="Soares infant brain lN1B"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lfamid BA; Site:1; Not
I; Site:2; Hind III; 1st strand cDNA was primed with a Not
I - oligo(dT) primer [5,
AACTGGAGATTCGCGCGCGGAGAAATTTTTTTTTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lfamid BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 126 a 121 c 135 g 103 t
ORIGIN

Query Match 6.6%; Score 92.2; DB 144; Length 485;
Best Local Similarity 57.4%; Pred. No. 1.6e-15;
Matches 166; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 376 ggttcgggttcacatcagcaaaacggctacatcctgacccaatacccaacgctcgttgc 435
DB 5 GGGTCGTGGGTTTATTGTCGGAAGATGGACTGATCGTGACAAATGCCACGTGGTGACC 64
QY 436 ggtatgggcagtcataaagtcctgtcacaacgacgaatataccgccaactcattc 495
DB 65 AACAGCAGCGGGTCAAGGTTGAGCTGAAGACGGTGCCACTTACGAAGGCCAAATCAAG 124
QY 496 ggttcggatgtccatcccatgctgcctctctgaaatcgacgcaacggaagagctacc 555
DB 125 GATGTGGATGAGAAACGACATCGGCACTCATCAAAATTGACCAACGAGGCAAGCTGCCT 184
QY 556 gtcgtcaaaatcgcaatcccaaaatttgaaacccgggggaatgggtcgctgcacatcgc 615
DB 185 GTCCTGCTGCTGGCGCTCTCAGAGCTGCGCGCGGAGAGTTGCTGTGTCGCCATCGGA 244
QY 616 ggcgccttcggctttgacaacagcgtgacgcgcgggcatcgtgtccgccca 664
DB 245 AGCCCGTTTTTCCCTTCAAAACACAGTCACACCGGGATCGTGAGCACCA 293

RESULT 9

AW660844 504 bp mRNA EST 14-JUL-2000
LOCUS 99933 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION
ACCESSION AW660844
VERSION AW660844.1 GI:7426671
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 504)
AUTHORS Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,
Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid
W.W. and Keele,J.W.
TITLE Design and use of four pooled tissue normalized cDNA libraries for
EST discovery in cattle
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred

Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@cocoa.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a

	SOURCE	FEATURES
	source	/organism="Bos taurus"
	1..529	
	Mouton/Judith	

```

BASE COUNT
ORIGIN
135 a 158 c 158 g 108 t
/note=Vector: pZLL; Site_1: SalI; Site_2: NotI; Poly A
was deleted from a NotI site"

```

Query Match	6.3%	Score 87.8;	DB 37;	Length 559;
Best Local Similarity	55.9%;	Pred. No. 3.1e-14;		
Matches 167; Conservative	0;	Mismatches 132;	Indels 0;	Gaps 0;

QY	366	attgaacttcgggttcgggttcctcatccagcaaaaacggctacatccttgaccaataccca	425
Db	168	AGTGGCCAGTGGGTCTGGGTTCATCGTGTCCGAAGATGGACTGATTGTGCAACACGCCCA	227
QY	426	cgtcgttcgccggtataggcagtcatacaaatcctgcctcaacgacgaacggaataaccgc	485
Db	228	CGTGGTGACCAACAACAGCATCGGGTCAAACTCGAGCTTAAGAACGGTGCACCTATGAGGC	287
QY	486	caaacctcattcgggttcggatgccatcccatgtcgccctcttgaaatcgacgcaaacgga	545
Db	288	CAAAAATCAAGACGTGGACGAGGAAGCGGCATATGGCCTTATCAAAATCGACCACCGGG	347
QY	546	agagctaccctgcgtcaaaatcggcaatcccaaaatttgaaaccggcggaatggatcgc	605
Db	348	AAAGCTGCCTGTGCTACTGTTGGCCGCTCCTCGGAGCTGCGGCCGGGAGAGTTCTGTGT	407
QY	606	tgcacatgcgcgccttcggctttgacaacagcgtgacccgcgggcacatcgtgtccgcaca	664
Db	408	CGCATATGGAAGCCCGTTTTCCTTCAAAACACGGTCACACCGGGATGTCAGCACCA	466

RESULT	13
LOCUS	AI903651/c
DEFINITION	AI903651 337 bp mRNA
ACCESSION	QV-BT032-190299-136 BT032 Homo sapiens cDNA, mRNA sequence.
VERSION	AI903651
KEYWORDS	AI903651.1 GI:6494038
SOURCE	EST
	human

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP
Brazil
Tel: +55-11-2704922

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP
Brazil
Tel: +55-11-2704922

Job time: 4478 sec

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE AUTHORS

1 (bases 1 to 288)
Hallier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman
M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston
R., Williamson, A., Wohlmann, P. and Wilson, R.
The WashU-Merck EST Project

TITLE JOURNAL COMMENT

Unpublished (1995)
Contact: Willson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

High quality sequence stops: 143 Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: similarity on wrong strand
Insert Length: 1530 Std Error: 0.00
Seq primer: M13Rpr1

High quality sequence stop: 143.

FEATURES source

Location/Qualifiers
1..288
/organism="Homo sapiens"
/db_xref="GDB:397142"
/db_xref="taxon:9606"
/clone="IMAGE:24795"
/clone.lib="Soares infant brain L1NB"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lfamid BA; Site.1: Not
I; Site.2: Hind III; 1st strand cDNA was primed with a Not
I - oligo(dT) primer [5',
AACTGGAGAAATTCGCGCGCAGGAAATTTTTTTTTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lfamid BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT
ORIGIN

68 a 71 c 75 g 61 t 13 others

Query Match 5.8%; Score 81.2; DB 146; Length 288;
Best Local Similarity 53.7%; Pred. No. 1.e-12;
Matches 152; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 374 tcggttcgggttcacatcagcaaaaacgggtacatcctgacaaatcccaacgtcgttg 433
Db 5 TGGGCTCGGTTTATTGTGCGAAGATGGACTGATCGGTGACAAATGCCACCGTGGTGA 64
QY 434 ccggtatgggcagatcaaaatcctcgtcaacgacaagcggaataatccgcgaactca 493
Db 65 CCAACAGACCGGGTCAAGTTGAGCTGSAAGACGGTGCCACTTACGAAGCCAAATCA 124
QY 494 tcggttcggatgtccaatccgatgtcgcctctctgaaaatcgacgcaacgagagctac 553
Db 125 AGGATGTGGATGAGAACGACACATCGCATCATCAAAATTGACCACCAGGCGCAAGCTNC 184
QY 554 ccgtcgtcaaaatcggaatccccaaaatttgaaccggcggaatgggtcgtcgtccatcg 613
Db 185 CTGTCCTGCTNCTTTGGCNCCTCCAGAGCTCGGCNGGAGAGTTCGTNGTCNCCATCG 244
QY 614 qcgcccttcgggtttgacaaacggtgacgcggcgatcgt 656
Db 245 GANCCCNCTTTCCCTTCAAAATCAGTANANCNCGGGNCTCGT 287

Search completed: March 22, 2001, 03:35:13

